

Nature's East–West exchange

In the knowledge that the scientific enterprise is genuinely international and in the belief that Mr Mikhail Gorbachev's speech to the United Nations on 3 December offers us all the promise of a different world, Nature makes the following announcement to and solicitation of its readers.

WHILE on the face of things there may seem little that the scientific community can do, except to keep on working, to assist or cement the recent transformation of East–West relations, that is not strictly the case.

First, there is ample evidence that the huge Soviet research community, hampered as it is by bureaucracy and shortages of equipment, would benefit from closer relations with the outside world. Individuals might thereby be helped to be more creative. Research laboratories and institutes could well be more productive and effective if their inevitable isolation were abated even modestly. There can be few in the West who would not acknowledge that these benefits must be mutual.

It is also plain that Soviet science has much to offer, even in these competitive times, to people working in the relatively well-equipped laboratories of the West. Those benefits would also be mutual.

In the past few years, public authorities in the Soviet Union and elsewhere have recognized the urgency of these needs. Opportunities for Soviet scientists to travel to the West have been enormously increased, and at the same time liberalized. Travel in the opposite direction has similarly been enlarged. Research collaborations have multiplied, if less quickly. But it is also clear that there is a long way to go before relations between the Soviet research community and those elsewhere are as intimate and constructive as they might be.

Nature therefore intends to take two practical steps to assist these processes.

- First, we shall do more than has been possible so far to draw attention to important trends in Soviet research, largely for the benefit of readers in the West. Talks with the Academy of Sciences of the USSR, as well as with individual scientists, encourage the belief that there is much that can usefully be done.

- Second, *Nature* will create an informal network, to be known as *Nature's East–West Exchange*, whose chief objective will be to put professional researchers with mutual interests in touch with one another. *Nature* will function as a post-box only, relying primarily on information voluntarily supplied by readers, but will also use its own network of advisers to identify potential respondents to particular enquiries.

It would be rash to guess how well and how generally this arrangement can be made to work, let alone to estimate its importance in the grander scheme of things. But a scheme for exchanging information and easily transportable materials which is entirely independent of government agencies and established academies may have

the special value of enabling informal and, with luck, speedy responses to enquiries.

None of this should be taken as diminishing the importance of the bilateral exchange schemes at present operated by national academies and other agencies, or the arrangements for collaborative research involving Soviet scientists and their counterparts elsewhere. It is also a fact that in some fields, high-energy physics or plasma physics for example, Soviet collaboration with others has always been close. An informal private scheme for the exchange of information may nevertheless usefully complement arrangements already in place.

Readers wishing to participate should send in writing their names, addresses, telephone and telefax numbers to one of the editorial offices in London, Munich, Tokyo or Washington. Readers are assured that this information will not be used for any other purpose.

Readers should indicate their interest in the scheme, which may take one or more of the following forms:

- A request for information or assistance, which should be as specific as possible.
- A statement of particular information and assistance that can be conveniently provided, and which is unlikely to be readily available in the East (or West), which should also be specified as fully as possible.
- An offer to help in identifying sources of information and assistance in particular fields of research, which should also be specified with some precision. ("High-energy physics" or "molecular genetics" may invite more requests for help than respondents would wish.)

It is not intended that the scope of the exchange should be limited, but it is expected that enquiries will at the outset be requests for information, perhaps for software developed specifically for basic research, but physical samples (say, DNA probes) may also be suitable for exchange. Respondents should be aware that it is for them, not *Nature*, to comply with whatever national legislation may be in force.

Particular care will be taken to ensure that Soviet scientists at Soviet universities and less well-known research institutes will be able to collaborate.

No charges should ordinarily be made for information or materials supplied under this exchange, although there may be circumstances when two participants agree that costs arising should be paid for.

Nature will publish reports on progress of the scheme at intervals of three months.

A list of people who have endorsed the general objectives of this scheme will be published next week. □



Europe in trouble

The impending trade dispute with the United States will confirm fears about the new Europe.

THE European Community seems well on the way to being an obstruction to international amity. That is the chief lesson that will be learned from the foolish dispute with the United States over the import of beef which was institutionalized last Sunday, the first day of the New Year. The most immediate consequence will be to sour the atmosphere in which some kind of compromise must be hammered out on the future of the international agreements that make the world's free trade an engine of beneficent economic change; after the collapse last December of the Montreal negotiations of an extended basis for the General Agreement on Tariffs and Trade (GATT), it was agreed that the participants have only until April to settle their differences.

But the more distant yet more serious consequences of the dispute will be to confirm the impression that the new Europe is bent on being a selfish and restrictive member of the international community, more concerned with its own interests than with the well-being of the whole world, which is what the critics of the European Community have always feared.

It is especially discreditable that this year's trade dispute should spring from what is represented as a scientific issue. Two years ago, the European Community took the view that beef cattle nurtured with the help of bovine growth hormones would yield beef that endangers the health of those who eat it. There appears to be no factual basis for this supposition, but every reason to suppose that the truth is that beef from cattle grown in such a way is identical with what might be called natural beef.

Certainly the European Commission's directive which, from 1 January, bans the sale of beef whose growth has been assisted by the administration of growth hormones is innocent of an objective test for telling now-tainted beef from that whose sale is to be allowed. Instead, those wishing to import beef into Europe have to obtain a certificate from suppliers elsewhere that their cattle have not been fed growth hormones to help them grow more quickly. And there seems no doubt that the motives for the new restrictions on the sale of beef grown with the help of hormones are simply the commission's wish to make European agriculture less efficient than it could be, and thus a less embarrassing charge on European budgets. Briefly, Europe has decided to ban the sale of beef grown with the help of growth hormones because it does not wish to pay European farmers even more than it does at present for the accumulation of unsaleable stocks of European beef. In a rational world, of course, the remedy would have been much simpler; Europe would have reformed its irrational policy on agriculture instead.

Response

The immediate response of the United States is understandable. With effect from 1 January, retaliatory tariffs will be applied to certain imports of agricultural produce of European origin.

The losers will be of two kinds — consumers in the United States and farmers in the European Community. But there is worse to come. For the European Community now threatens to retaliate against the retaliation instituted by the United States. There is every likelihood that European agriculture ministers, who have many better things to do, will spend their energies in the coming weeks devising lists of agricultural products from the United States to which further restrictions may be applied. The losers, then, will be European consumers, who will have to pay higher prices for goods to which they have become accustomed.

But there is every likelihood that mutual retaliation will escalate, causing people on both sides of the Atlantic to become impoverished and creating a climate in which sensible agreements on the future pattern of free trade cannot be reached.

The prospect of a trade war in agricultural produce between Europe and the United States is, sadly, not the only ominous sign of the way in which modern Europe is tending. In the past few months, the European Commission has responded to the pleadings of European manufacturers by deciding unilaterally that Japanese exports of copying machines and electronic printers are too cheap, and has slapped restrictions on them as well. Japan has rightly decided to appeal against these arbitrary decisions, as GATT procedures allow, but that is a time-consuming and expensive business whose outcome must be uncertain.

Before that corrosive issue can be settled, there will no doubt be other European manufacturers pleading successfully for protection from imports from elsewhere which they consider are too cheap, meaning simply that the prices at which goods are offered for sale are inconveniently and unprofitably low. It may not be long before the manufacturers of civil aircraft, or computing machinery, lodge successful claims that they, too, should be protected from competition from elsewhere. While, in the short run, European manufacturers may enjoy the benefits of markets in which they are the sole suppliers, Europe as a whole will lose both the benefits of competitive prices and of underlying advanced technology.

Gloomy outlook

The outlook is all the more gloomy because the United States, the most likely target of European protectionism, is now equipped with domestic trade legislation that makes retaliation almost a statutory requirement. While the Reagan administration has steadfastly, over eight years, resisted attempts by the US Congress to protect particular industries from competition from elsewhere, it was forced last year to concede a provision that industries with cause to believe that their competitiveness in export markets is thwarted by foreign restrictions should have a right to demand retaliation. Already there is an awesome queue of claimants. It will be especially difficult for a new administration, needing to plead with Congress on the federal deficit, to keep these domestic demands at bay. In short, there is a serious prospect that the attempt to negotiate a new basis for an extension of GATT's terms of reference will be aborted by a trans-Atlantic trade war extending well beyond the field of agriculture.

If that is what happens in the months ahead, Europe will have to bear the chief responsibility. But, sadly, Europe is not a single entity, but rather a collection of twelve governments with different interests, most of which are narrow interests. Decisions made in Brussels by the European Commission ultimately reflect the way in which particular governments are able successfully to plead their own causes. That is the reason why the common market fully specified by the Treaty of Rome has not yet come about — and why there remain serious dangers that the creation of the single market now planned for 1992 may itself be aborted by strategems not yet conceived. The same narrow view of European affairs offers an explanation as to why it has been so difficult to move towards a rationalization of the agricultural policies that are Europe's self-inflicted burden.

The next few years unfortunately offer narrow-minded European governments too much opportunity for insisting that further moves towards internal free trade must be accompanied by economic protection from outside competition. The ironical result of that will be a Europe which will be self-contained and inward-looking — one in which Europe's founding fathers would prefer not to live. □

International expedition studies Arctic ozone

- Differences in Arctic and Antarctic depletions
- Specially designed aircraft to be used

Washington

AN international expedition to study stratospheric ozone concentrations in the Arctic begins this week in Stavanger, Norway. Led by the United States, the six-week expedition will employ balloons, satellites, high-altitude aircraft and an international data network to obtain a clearer picture of how Arctic ozone concentrations are affected by the unique conditions at the pole.

Arctic stratospheric ozone concentration exhibits some seasonal variation, but nothing like the severe declines observed over Antarctica. Nevertheless there are indications that similar processes are at work in the Arctic stratosphere, leading to a measurable depletion of ozone. More puzzling, according to Daniel Albritton of the National Oceanic and Atmospheric Administration (NOAA) Aeronomy Laboratory in Boulder, Colorado, is a discrepancy between observed ozone depletion and predicted levels based on models that account for Antarctic depletions, suggesting that some other mechanism is at work in the Arctic.

Two specially designed aircraft will be used to collect data on gas concentrations in the stratosphere. A National Aeronautics and Space Administration (NASA)

ER-2 aircraft, a modified U2 reconnaissance aircraft, will carry instruments as high as 20 km for seven-hour flights up to 80 degrees North. A DC-8 will operate during the same period, its greater range allowing it to survey the polar vortex. NASA's Robert Watson, chief program officer for the expedition, says that ozone measurements from the Nimbus 7 satellite will help to develop flight plans for the aircraft.

Of particular interest is how chlorine compounds — primarily from chlorofluorocarbons — react with ice crystals in polar stratospheric clouds to form free chlorine which destroys ozone. Because the Arctic is on average warmer than the Antarctic, there are far fewer of these ice-containing clouds. Albritton says this may partly explain why the decrease in stratospheric ozone is not nearly as large as in the south. Albritton is hoping for a cold polar winter so there will be lots of polar stratospheric clouds to test the ice crystal hypothesis.

NASA, NOAA and the National Science Foundation together are contributing some \$10 million to the expedition. Additional support, comes from Norway, the United Kingdom, West Germany and the Soviet Union.

Joseph Palca

Computer smugglers too quick

Washington

LAST week's arrest in Miami of a Dutch man accused of attempting to smuggle a Vax 8800 computer to Bulgaria highlights the difficulties the United States is experiencing in stopping advanced computer technology from reaching eastern-bloc countries. The difficulties will get worse, according to a new report* on export controls and computer technology released this week by the US National Research Council, as computers become ever smaller and more powerful and international computer networks flourish.

The nations of the Coordinating Committee for Multilateral Export Control (CoCom) have managed to maintain a five-to-ten-year lead in computer technology over the USSR and eastern-bloc countries, claims the report, but far more "focused and flexible" controls will be needed in the future. The biggest problem is that "computer technologies advance much faster than most bureaucratic processes".

A clumsy bureaucracy will end up stifling industry if it does not distinguish

between products that can be controlled and those that cannot — such as computers available from foreign sources and computers that are so cheap or small that their trade cannot be effectively halted. "More timely, rapid and expert" reviews are the answer, according to the report.

To help stop the industrializing Asian nations of Singapore, Hong Kong, Taiwan and South Korea from supplying their own increasingly sophisticated hardware to the eastern bloc, the report urges their integration into the CoCom programme.

The rapid development of computer networks has provided an easy conduit for the export of software. The report calls for "further study" to control transborder flow via computer networks.

Software is intrinsically difficult to protect and most should be freely traded, according to the report, with tight controls reserved for military software and for software tools that could help create military software.

Alun Anderson

*Global Trends in Computer Technology and Their Impact on Export Control, published by the National Academy Press, 30 December 1988.

Ecologist murdered

Washington

LAST week's murder of the prominent Brazilian environmentalist and union organizer, Francisco Mendes Filho, at his home in Xapuri, in Acre province, is expected to increase pressure on the World Bank and Inter-American Development Bank to demand environmental safeguards before they lend money for large-scale development projects. Through his efforts to organize and protect poor rubber-tapping families Mendes, 44, inadvertently found himself allied with ecologists who wanted to stop the destruction of the rain forest for cattle ranching. His murder, allegedly at the hands of local land barons, has provoked an international outcry and a campaign in Brazil to stop foreign loans until justice has been done. Last year, Mendes received the Global 500 award from the United Nations for his efforts to save the rain forest.

A.A.

Shuttle complacency

Washington

ALTHOUGH the US space shuttle has successfully resumed flights, the final report of a special National Research Council panel evaluating the redesign of the shuttle's solid rocket booster concludes that there are still problems with the booster.

In addition to specific, minor problems that cropped up on the two recent shuttle flights, the final report, transmitted in a 21 December letter from panel chairman H. Guyford Stever to National Aeronautics and Space Administration head James Fletcher, suggests there is still work to be done. The report warns that budget, time and complacency will push NASA towards a relaxation of quality control. Although the "relaxation" may be justified, the report urges that it be a planned, deliberate process, and calls for continuous system performance checks to establish a data base against which the shuttle's future performance can be judged.

J.P.

Animals' legal rights

Washington

A GROUP of animal rights organizations, led by the Animal Legal Defense Fund, has filed suit against three US government agencies. The group is accusing the Department of Agriculture (USDA), the Department of Health and Human Services, and the Office of Management and Budget of "unreasonably delaying" the establishment of regulations to protect animals used in research which were stipulated by the 1985 amendments to the Animal Welfare Act. The amendments would require researchers to anaesthetize animals subjected to pain, and to provide dogs and primates with living environments conducive to their well-being. The USDA says it is still sifting through the roughly 8,000 comments it has received.

C.E.

'Virus-proof' computer security system

Boston

CAN computer operating be made safe from invading viruses, worms and determined hackers? Two computer scientists think the answer is "yes", and they have developed a computer security system which they claim is invulnerable to computer viruses such as the one that recently struck networks across the country (see *Nature* 336, 97 and 301; 10 and 24 November 1988).

The system, designed by Michael Rabin, a computer scientist at Harvard University, and his former student Douglas Tygar, now assistant professor in the School of Computer Science at Carnegie-Mellon University, is an adaptable set of programs designed to regulate access to information on computer systems.

The system, called ITOSS (Integrated Toolkit for Operating System Security), is the culmination of a four-year effort by Rabin and Tygar. Among its features is a battery of built-in 'sentinels' which can be employed either to warn away users from high-security data, or to record every access to sensitive data files without users'

knowledge. A 'fingerprinting' algorithm can also alert system operators to any changes made to programs on the system. Included in the program's design are a 'minimum security' protection for all files against 'trojan horse' programs that might come into the system illicitly, hidden in innocuous messages, and a 'maximum security' feature which requires a specified number of system operators each to input commands in order to make any far-reaching changes in the system.

New systems to protect computers are likely to be necessary if computer operators are to gain the upper hand over hackers. Last month, officials at Lawrence Livermore Laboratory gave up their battle with a hacker who had repeatedly infiltrated low-level security files on the laboratory's computer system. Instead, they left an invitation for the hacker to make direct contact with the laboratory's computer security manager. The hacker accepted the invitation and has now promised to stop his attacks. An FBI investigation is nevertheless under way.

Seth Shulman

Increase for West German ocean research

Munich

A SIZEABLE increase in support for ocean and polar research in West Germany was recommended last week by the influential Wissenschaftsrat, an independent science advisory board. The board recommended that an additional DM240 million should be spent over the next few years, with the money to come from both the federal and *Länder* governments.

The largest new investment recommended is for a DM75 million research centre in Hamburg which would bring together researchers from the University of Hamburg, the Max Planck Institute for Meteorology, the German Climate Computing Centre and the Federal Biological Research Establishment at Heligoland.

Research should be concentrated, the board says, on interactions between ocean and atmosphere as well as on the study of

structure and function of marine ecosystems. The four institutions involved, which are funded out of as many different purses, would have to agree on how to carry out such an integration. Two new institutes are recommended for the Oldenburg/Wilhelmshaven area, one to study the chemistry and biology of the ocean and the other to study coastal wetlands. Together they would cost DM40 million.

The Wissenschaftsrat also praised the Alfred Wegener Institute for Polar and Ocean Research in its first eight years as a *Grossforschungseinrichtung* (large research establishment). The leader of the Bremen institute, Gotthilf Hempel, wants to broaden its focus to include Arctic as well as Antarctic research and to include even more international cooperation. Hempel is chairman of the Arctic Ocean Sciences Board.

Steven Dickman

£520 million plans for UK science park

London

EVERY British university is trying to broaden its funding base and to forge links with industry. With that aim, a business park with a science centre at its core is planned for a 500-acre site near the University of Bristol. Details of the plans were released on 19 December.

The £520-million venture will include facilities for research, development and manufacturing, with space for 22 companies. These will have access to the 150 researchers based in the science centre who will come from the Universities of

Bristol and Bath and the Polytechnic of Bristol. Each institution has a 5 per cent stake, which is paid for in kind, with the provision of researchers and support staff.

A public inquiry must be held before the venture is approved because the site is on land released from environmental protection regulations. But Don Carleton, of the University of Bristol and director of the development company proposing the plans, says the planning authorities support the plans. He expects construction to begin next summer, and the park to create 14,000 new jobs.

Christine McGourty

Tanks to hold more than fuel

Washington

THE US space shuttle's giant external fuel tank will some day carry scientific experiments of its own, according to a memorandum of agreement signed last week between the National Aeronautics and Space Administration (NASA) and the Space Phoenix Program of the University Corporation for Atmospheric Research.

In his 1988 Commercial Space Initiative, President Ronald Reagan offered the external fuel tanks to anyone who could think of something interesting to do with them (see *Nature* 331, 550; 1988). At present, the tanks are unceremoniously dumped into the ocean after providing the propellants to boost the shuttle into orbit.

There is an unpressurized, 5,000-cubic-foot space between the fuel and oxidizer tanks that can hold a pallet containing scientific instruments. According to Peter Riva, a spokesman for the Space Phoenix Program, that space can easily accommodate relatively large, low-technology experiments that can be loaded just before launch. Some experiments now being considered are gas releases into the upper atmosphere and detailed studies of the thermosphere.

Initial plans call for using the external tank during sub-orbital flights, but ultimately the Space Phoenix Program would like to find a way to place the tanks in more permanent, slow-decay orbits. There is no fixed timetable for when the external tanks will be used to carry experiments, but Riva expects the first flight will be sometime in 1991.

Joseph Palca

Neuroscience Center

Boston

THE Harvard-affiliated Massachusetts General Hospital has opened a new \$18 million interdisciplinary Neuroscience Center. The centre, made up of 11 separate laboratories and with a staff of more than a hundred scientists, students, fellows and technicians, is financed by a \$6.4 million grant from E.R. Squibb and Sons, a \$5 million grant from the National Institutes of Health, and by a number of other private sources.

Dr Joseph Martin, professor of neurology at Harvard Medical School, will direct the new centre. Martin stresses the interdisciplinary approach of the centre, saying that chronic neurodegenerative diseases might best be tackled by "understanding their causes at the genetic level" rather than a "shotgun" approach of focusing only on particular symptoms of individual diseases. The centre will house the hospital's neurogenetics laboratory where molecular genetic techniques are being used to study Huntington's disease and Alzheimer's disease.

Seth Shulman

20th-century smugglers use biotechnology

Sydney

IN a cloak and dagger operation that could have been lifted from the pages of a spy thriller, a group of wily Australians have combined the traditional trickery of smuggling with high-technology embryo implantation techniques to try to break South Africa's near-monopoly of the \$200 million market for mohair wool.

Mohair wool comes from angora goats, the finest of which are bred in South Africa. Exporting them without the permission of the South African National Angora Breeding Society is banned. But in 1986, three Australian farmers raised A\$600,000 from Australian goat breeders to try to lay their hands on a herd.

In South Africa they succeeded in buying 269 top quality Angora and Boer (meat) goats. Waiting until Christmas Eve, when they guessed official diligence would be low, they loaded the herd into trucks and set off at night for a 47-hour trip to the border with Zimbabwe. One farmer flew ahead of the convoy in an aeroplane, smoothing out paperwork at more than 20 checkpoints along the way. At the border, customs inspectors were persuaded the trucks contained sheep rather than goats, a move one of the men involved in the trip agreed was "against the spirit of the South African angora industry regulations".

In Zimbabwe, the Australians bred the goats and collected almost 400 embryos which were frozen and flown to the Australian Government offshore quarantine station on the Cocos Islands in the Indian Ocean. The plan was to implant the embryos into surrogate mothers on the Cocos Islands and keep the kids in quarantine. But the original partnership broke up and the frozen embryos were instead sold to Embryotech, a Western Australian company. Embryotech transported the frozen embryos to New Zealand, which has already built up a herd of 74 Angora kids and 24 Boer kids.

With multiple ovulation and embryo-transfer technology, the number of offspring can be rapidly increased. Once seven years of quarantine are up — essential to ensure the herd is free from the scrapie disease agent — Embryotech has just announced that it will sell all the offspring to Australian farmers, giving them a chance to cash in on the lucrative angora wool market.

Three other angora herds have also been 'imported' to New Zealand from South Africa. In 1987 South Africa introduced legislation that allows the government to confiscate the farm of anyone collaborating in illegal exports.

Tania Ewing

A certain uncertainty awaits in Washington's new year

Washington

CELESTIAL mechanics guarantees that in August this year the Voyager space probe will make a historic encounter with the outer planet Neptune. But apart from that, nearly every other scientific endeavour of 1989 is tinged by an uncertainty not felt for 8 years: from 20 January a new man will be in the White House, and although President-elect George Bush is of the same political cloth as his predecessor, there are already indications that new priorities can be expected.

During his campaign for election, Bush promised to raise the status of the White House science adviser to the rank of presidential assistant, but by the year's end when — nearly — all cabinet appointments had been announced there was still no word about who will fill that new position. This may be an indication that science will not in fact receive its promised new importance, but it could just as easily be evidence that Bush has found — as have others before him — that the political and bureaucratic complexities of bringing in a new team can foil the best intentions.

Adding to the uncertainty are a spate of new faces in important congressional appropriations committees. This month, Congress will receive the 1990 budget blueprint prepared by the Reagan administration, but that will certainly be modified by both the Bush administration and Congress before the start of the new fiscal year.

In addition to establishing budget priorities that could spell the end of big science projects such as the Superconducting Super Collider — now scheduled to make Waxahachie, Texas the mecca of the high energy physics world — and the National Aeronautics and Space Administration (NASA)'s Space Station, Congress will also continue to grapple with the issue of scientific misconduct.

The arcane question of the specificity of the Bet-1 reagent may be the central theme of hearings scheduled for early in the year, as a powerful committee led by Representative John Dingell (Democrat, Michigan) resumes its inquiry into a paper appearing in *Cell* co-authored by, among others, Nobel laureate David Baltimore, that has raised complicated issues of how the propriety of scientific data should be investigated. Another committee chaired by Representative Ted Weiss (Democrat, New York), will continue hearings probing conflict-of-interest issues for scientists who work at academic institutions but are supported by private industry.

For NASA, 1988 was a good year, and 1989 promises to be even better. The space shuttle returned to service in Sep-

tember, and a second successful flight followed. Three major scientific missions are scheduled for launch this year; the Magellan Venus radar mapper, the Galileo mission to Jupiter and the Hubble Space Telescope. If, however, the Bush administration decides not to go forward with the space station, NASA will be searching for a new identity, as military and commercial space activities have already been moved elsewhere.

NASA will certainly have a role studying the environment, a topic that will occupy centre stage in 1989. Last summer's drought and heat wave convinced many that greenhouse warming was worth serious study, and finding ways to limit the production of greenhouse gases, including renewed consideration of nuclear power, will be a constant theme.

If new commercial nuclear reactors are still some way off, the Department of Energy must nevertheless finalize plans for a new defence production reactor to restore US capacity to produce tritium and plutonium for nuclear weapons. By the end of 1988, all US reactors capable of producing tritium had been shut down, and a constant stream of newly discovered threats to safe operation of the ageing reactors threatened any early restart.

The National Science Foundation (NSF) is still hoping to see its budget double by 1992, an increase promised by the last administration, but probably out of reach. NSF has moved ahead by providing money for 11 new science and technology centres, and more may be authorized in 1989.

Coping with AIDS will continue to be a major preoccupation of the US Public Health Service and its subordinate National Institutes of Health (NIH) in Bethesda and Centers for Disease Control in Atlanta. Also in Bethesda, the NIH human genome office, led by James Watson, will get into full swing next year, and W. French Anderson, along with colleague Steven Rosenberg, should begin the first-ever experiment involving the introduction of genetically engineered cells into human subjects.

Biomedical research will continue to face stiff challenges from two well-organized lobbies. Animal rights activists will continue to push for restrictions on the use of animals in research, and anti-abortion groups will seek to prohibit the use in research of fetal tissue from induced abortions. But Bush's choice of Louis Sullivan for Secretary of Health and Human Services has at once outraged anti-abortion lobbies and proved that what was past is not necessarily prelude.

Joseph Palca

Indian agreement with China sets the scene for 1989

New Delhi

In India, 1988 was seen out with the signing of a new agreement with China on cooperation in science and technology. After having having been cut off from one another for 30 years, scientists of the two countries will now cooperate "on the basis of equality and mutual benefit".

Set alongside increasing technological ties with the Soviet Union — including a promise to aid construction of two nuclear reactors on India's east coast — the agreement may signal a trend for Indian scientists to forge closer ties with colleagues in the communist world. Indo-US cooperation reached a low point in 1988 with dis-

putes over patents and protection of intellectual patent property rights.

The five-year science pact with China is one of the three agreements signed during Prime Minister Rajiv Gandhi's pre-Christmas visit to Beijing. The agreement will provide a framework within which Indian and Chinese institutions can conclude protocols and research contracts. Agriculture and industry will take priority but other forms of technological co-operation may be mutually agreed upon. The agreement also provides for exchange of scientists, documents and journals, and holding of seminars and workshops. From a broader political viewpoint, this agree-

ment and two others on culture and trade are seen as a prelude to an accord on the contentious border issue.

Earlier in 1988, India's success with its first remote-sensing satellite (IRS-1) was offset by the failure of the ASLV-3 launcher and the trouble with communications satellite INSAT 1-C.

India's ambitious space programme hinges on the next flight of ASLV-3 and the May 1989 launching of INSAT-1D. The heavy launch vehicle programme could even be terminated if India decides to concentrate instead on the hyperplane, the Defence Department idea for a spaceplane that collects oxygen fuel in flight. Soviet help has been sought.

Indian scientists scored a success with their first low-cost parallel-processing computer and by predicting accurately the behaviour of the monsoon for the first time since the creation of the meteorological department. The government will now launch a three-year project to build a supercomputer.

In the medical field, the birth control vaccine developed by the National Institute of Immunology will enter phase-two trials this year. Another vaccine for sterilizing stray dogs and unproductive cattle has been developed by the same institute and is now for sale.

Materials science, biotechnology and energy are key areas for funding in 1989. The first wave-power plant is being built on the southern coast and a plant is to be set up in Gujarat to extract gas from coal without mining. One nuclear plant will be commissioned in 1989 and 20 more will be built over the next 11 years. The Department of Atomic Energy will have no problems with funds but will find it increasingly difficult to overcome opposition from environmentalists.

On the science policy front, Gandhi has still not acted on his advisers' suggestion for a national commission to integrate science and technology with economic planning. Two controversial decisions on liberalizing seed import and permitting entry of Pepsico in the beverage market have been blamed on the lack of a well-defined technology import policy.

Gandhi has asked scientists to help solve India's problems through technology missions. Telecommunications has provided the first success. India has built its own electronic telephone exchange for cities and the mission intends to supply one exchange every day to rural areas. Over the next few years, backing will be given to missions for drinking water, immunization, cattle improvement and mosquito control. But as yet there is no mission to relieve the sufferings of the thousands of Bhopal gas victims who observed the fourth anniversary of the tragedy on 3 December by burning effigies in front of the Union Carbide factory.

K.S. Jayaraman

Still no hope of a coherent Australian science policy

Sydney

AT the end of 1988, Australian scientists finally found their voices. Dissatisfaction over reduced funding and the constant calls by the government to make science more commercial angered scientists sufficiently for them to reject the label of "wimps" given by Minister of Science Barry Jones and come out of their laboratories and protest.

Hopes had been raised that a coherent long-term science policy would emerge from the reports of a government inter-departmental committee (IDC) set up to look into funding and policy needs and from the Smith committee investigating university research requirements, even though the former contains no members from the scientific community.

But scientists seem likely to be further disappointed. The IDC report was presented to the cabinet in December and, although it has not yet been made public, details of its recommendations are known. The report acknowledges the problems of Australian science: poor funding, increasing student-to-staff ratios and the reduction in graduate students, but warns that any general topping-up of funds might be interpreted as "too little too late . . . The Government (would be seen to be) wavering in its policy directions".

In response to criticism that the administration of science is too fragmented, the report recommends a permanent committee of senior bureaucrats reporting to the Minister of Science to co-ordinate policy and funding. A "distinguished scientist" would also be appointed for a two-year term as Chief Science Adviser to the minister.

The IDC report allocates A\$20 million to meet the most urgent needs for labora-

tory equipment in the year ending June 1989. But the funds cannot be given out until the Smith committee reports in late March 1989. The amount is much less than that requested by scientists. CSIRO alone requested an extra A\$75 million for next year.

The government seems relentless. A spokesperson for Senator John Button, the Minister for Industry, Technology and Commerce said "we are trying to force scientists to get funding from the private sector. Historically scientists go away and set their own priorities. They will just have to get used to living in the real world."

But big business may not be interested. Dr Max Whitton, chief of entomology at CSIRO, said "science cannot attract commercial funding if there isn't the viable operating base to attract partners". Professor David Penington, vice chancellor of the University of Melbourne, agrees. "Industry keeps its distance from us rather than developing a vigorous partnership. The idea that this could be resolved through a directive from Canberra is nonsense. Incentives and deregulation are needed."

Australia's scientists are fearful that in 1989 the downward spiral for research funding will continue. The Federation of Australian Scientific and Technological Societies (FASTS) with 60,000 members is determined to have science funding put on the political agenda, particularly with a federal election likely at the end of the year. Professor Frank Larkins, president of FASTS, said: "Fundamental science is the foundation of applied science."

"If the government wants science and technology to pay economic dividends, we need increased funding and a five- to ten-year policy."

Tania Ewing

Too many students, too little cash in German universities

Munich

1988 ended on a note of uncertainty in West Germany with students demonstrating at 40 universities to protest against overcrowding and lack of funds. Overcrowding is at record levels, with at least twice as many students registered as the universities are designed to handle.

A new DM2,100-million government 'emergency programme' announced in mid-December will help, but will make a real difference in only a few faculties. And even if the height of the *Studentenberg* begins to come down in 1989, many professors will continue to be hard-pressed to find time for their research.

University researchers have also been hit by a budget squeeze. The Deutsche Forschungsgemeinschaft (DFG) — the main university research funding body — found it could fund less than half the applications in its priority programmes (*Schwerpunktprogramme*). Overall, DFG turned down more grant applications than ever before, raising worries that young people will be forced out of research careers. But DFG takes heart from the better quality of grant proposals.

Researchers at the Max Planck Society are better off. A modest increase in the 1989 budget comes on top of Nobel prizes in 1988 for three of its researchers: Hartmut Michel, Johann Deisenhofer and Robert Huber. Michel has since moved to Texas, but Deisenhofer and Huber remain at Max Planck Institutes.

The Large Research Establishments (*Grossforschungseinrichtungen*, GFE), like the Max Planck Institutes, do not have to worry about the annoyance of too many students. But both associations have promised to cooperate more with the universities, in keeping with the recommendations released in May by the science council (*Wissenschaftsrat*). The council fears an increasing division between universities and research institutes.

In 1989, GFE are likely to place greater emphasis on health research now that Harald zur Hausen of the German Cancer Research Centre is chairman. They have also announced their intention to work more closely with researchers in other European countries.

International cooperation is also the theme for West Germany's burgeoning space programme. The Research and Technology Ministry (BMFT) will increase its investment by 11.8 per cent to more than DM1,300 million in 1989, keeping pace with France and the other members of the European Space Agency. BMFT has stepped up its efforts on its second space laboratory, to be launched by the US shuttle in late 1991. West German groups

have also sent payloads on a Chinese rocket and on the Soviet probe to Phobos.

Cooperation may come harder in areas where ethics are involved. Differences began to appear in 1988 between West Germany and other European Community members about research on human embryos and the sequencing of the human genome. Both are emotive issues that will command more attention in 1989.

AIDS was less of a public issue in the second half of 1988, but the liberal policy of the West German government towards AIDS carriers may be in for some changes in 1989 under a new and untried Health Minister, Ursula Lehr, who replaced the veteran Rita Süßmuth in November. Peter Gauweiler, a hard-line conservative, lost the AIDS portfolio in a Bavarian state

government reshuffle. On the research front, a modest BMFT programme to establish centres of excellence for AIDS research has yet to bear fruit.

1988 began and ended with nuclear 'scandals' of different kinds. In January, hundreds of mislabelled barrels of nuclear waste turned up unexpectedly at nuclear waste storage sites across West Germany. In December, public trust was dealt another blow when a reactor mishap at the Biblis nuclear plant near Frankfurt was found to have been covered up. Nevertheless, Chancellor Helmut Kohl's centre-right coalition government held fast to its nuclear programme, just as it had after Chernobyl.

1989 will find legislators watching the clock as the Kohl government's second term winds down. The next elections for the Bundestag (parliament) are not due until 1990 but five Landtag (state parliament) elections in 1989 will set the stage for the big campaign. **Steven Dickman**

Return to the good old days of plenty for French researchers

Paris

FRENCH scientists began 1988 facing cut-backs in basic research spending and with instructions to begin paying homage to industry. That, said President Jacques Chirac and science minister Jacques Valade, was the way to bring France into a united European market in 1993, creating the opportunity to compete with Japanese and US technology. But elections in the spring brought a change of government and, with it, hopes of a return to the good old days. Radical changes were made by the incoming socialist government, but it was a long time before a coherent policy was formulated.

Minister for Research and Technology Hubert Curien's first move was to rectify "flagrant deficiencies" in research funding. A \$195 million handout was given in June, doctoral bursaries were increased and 150 new research posts created. In the autumn Curien finally gave substance to President François Mitterrand's promises with an extra \$484 million increase in his 1989 budget, a 7.5 per cent increase over 1988 and way ahead of inflation.

With the new government settling in during the summer, nothing new was done in the fight against AIDS, even though France continued to rank highest in Europe for the number of AIDS sufferers. Finally, in November, Curien gave researchers \$24 million, saying that if it was not enough they would get more, while Health Minister Claude Evén launched a new \$8 million AIDS information campaign. 1989 will show whether it succeeds in converting Roman Catholics to the condom, outlawed by the Pope.

The French space industry, major shareholder in the European Space Agency (ESA), continues to boom. Sixteen satellites have been launched in the past 15 months. Ariane 4, the new workhorse heavy-lift launcher, also successfully completed its maiden flight in 1988. Ariane-space enters the new year with \$2,240 million worth of orders for the launch of 36 satellites. Now all eyes are on Ariane 5, which will be used to send the Hermes shuttle on its way to Europe's Columbus contribution to the US space station, also finally approved in 1988. Work has started at ESA's Kourou launch site to build a new launch complex for Ariane 5, due to fly in 1995, to the chagrin of British space scientists whose HOTOL shuttle is still looking for a sponsor.

Reform, the unlucky buzz-word of the last government, has been replaced by 'evaluate'. The high levels of public spending needed to realize the new government's declaration of research and education as priorities will need careful housekeeping. Curien has introduced a new watch-dog committee to oversee the cost-effectiveness of research. 1989 will also see Jacques Benveniste's research on the 'molecular memory' of water scrutinized as part of a routine, but major 'in-house' review of his laboratory.

Besides 'evaluate', the government's will is to 'co-ordinate'. No major policy decision has been taken without an in-depth report with the emphasis on inter-ministerial co-operation. In France, 1989 will be the year of committees, but only time will tell if their designs produce horses or camels. **Peter Coles**

Japanese science looking to the West and thinking big

Tokyo

In Japan, 1989 begins with increasing concern that the Recruit stock market scandal, which has already forced the resignation of Finance Minister Kiichi Miyazawa and top leaders in industry, will sweep away others in high places. Further revelations of influence peddling in the political world may mar a year in which Japan wanted to 'internationalize' its image, not least in science and technology.

1989 will be an important year for 'big science'. The Ministry of Education, Science and Culture (MESC) will make available several tens of millions of dollars to begin development of the world's largest helical fusion device at a new institute in Gifu Prefecture. And the Science and Technology Agency will pour similar sums into development of the world's largest synchrotron which is expected to be sited at a "technopolis" in Hyogo Prefecture, near Osaka. Both machines are scheduled to begin operation in the mid 1990s.

Japan will continue its modest but successful space programme with the launch of three satellites. The National Space Development Agency will put up a meteorological satellite (GMS-4) and a marine observation satellite (MOS 1-b). In February, the Institute of Space and Astronautical Science (ISAS) will launch a satellite (EXOS-D) into the 'bow wave' of the magnetosphere above the polar regions of the Earth to observe the formation of aurorae.

Eyes will be on the Space Activities Commission to see if (as is expected) it will grant a request from ISAS to increase the size of its rockets. For the past 20 years, ISAS rockets have been limited to a tail diameter of 141 cm under an agreement between the Science and Technology Agency and MESC. But larger rockets are needed to fulfil ambitious plans to send probes to the planets and beyond.

The first batches of young Western scientists invited under new postdoctoral fellowship schemes will begin arriving in Japan in the next few months. The government fellowships were introduced in the face of US pressure on Japan to open up its research system to Western (in particular US) scientists, and became available shortly after signing of the US-Japan Science and Technology Agreement in Toronto in June. But the schemes are suffering from a lack of applicants and it is uncertain if they can be expanded as planned in fiscal year 1989.

Japan's Human Frontier Science Program, which will be set up as a foundation in Europe in October, may fare better. A 'pilot' programme of five grants offered by

the Ministry of International Trade and Industry (MITI) last summer received a flood of applicants from scientists of the summit nations.

The grants, for research on higher-order brain functions or molecular-level recognition and response, are open to scientists in institutions located in three or more summit nations, one of which must be Japan. When the foundation opens in Europe, the grant programme will be expanded and the foundation will also award postdoctoral fellowships and stage international workshops.

It remains to be seen how much control the Japanese government will maintain over the foundation. MITI projects have tended to suffer from inflexibility; much will depend on the make-up and nature of governing bodies to be established for the programme.

Speedy internationalization will be seen in the private sector. In May a consortium of UK, US and Japanese companies headed by Britain's Cable and Wireless and C. Itoh of Japan will begin international telecommunication services in competition with Japan's Kokusai Denshin Denwa (KDD). The consortium will also begin laying a trans-Pacific optical fibre cable to compete with a new KDD cable that will come into operation in March.

Research and development will begin to move abroad with the opening of an institute sponsored by electronics giant NEC at Princeton in May. The institute is expected to set a trend, as is a three-year joint agreement between Hitachi and Texas Instruments to develop 16-megabit dynamic random access memory chips.

The Recruit scandal is thus all the more damaging because it comes at a time when Japan is trying to improve its image overseas. The Recruit saga reveals that 'money politics' still rules Japan and little has improved since the Lockheed scandal of the 1970s which brought down former prime minister Kakue Tanaka. Many questions remain unanswered, particularly regarding the involvement of former prime minister Yasuhiro Nakasone in the affair. And although Prime Minister Noboru Takeshita is now riding high because of the successful passage of a major tax reform bill, he is not immune from attack.

Takeshita's strategy will be to try to put Recruit behind him by launching his pet 'home town' plan for the development of the regions and decentralization of Tokyo. But the popularity of his administration has plummeted and he will have much to do if the party is to choose him to continue as prime minister in the autumn.

David Swinbanks

Glasnost doing perestroika in

London

DURING 1988, the Soviet Union had to come to grips with *glasnost*. From the February fire at the Academy of Sciences library in Leningrad, through the post-humous revelations of Valerii Legasov about the Chernobyl disaster, and up to questions as to why, after so much research investment in earthquake prediction, Spitak and Leninakan were unprepared, people in authority have had to accustom themselves to the unfamiliar sensation of being publicly questioned.

In science and technology, passing responsibility for failures back and forth between the research and production sectors will no longer do. The All-Union Academy of Sciences has completed its replacement of superannuated personnel (and elected Andrei Sakharov to its presidency), but has had its own scandal: the destruction of a major archaeological site for the building of its Crimean rest home. There have been major reassessments of history, many posthumous rehabilitations and 'rediscoveries' of disgraced party leaders, scientists, and scholars.

Yet *perestroika* (restructuring) is running slower than *glasnost*. Self-financing for industry and family leasing for agriculture have yet to show their economic results, and some publicized reforms have failed to materialize. Many managers still seem beset by the 'gigantomania' of the Brezhnev years. The most promising development of the year — an education reform that, it is promised, will at last allow excellence to flourish — was announced only in its closing days.

A major feature of 1988 has been the increase of 'national' demands throughout the 15 constituent republics of the Soviet Union. Some of the demands — such as the development of scientific vocabularies for the 'national' languages based on 'universal' rather than Russian roots, or the establishment of a separate Academy of Sciences for the Russian Federal Republic — should not prove too difficult to satisfy. Nor should some demands of environmentalists — such as the campaign against the chemical works in Erevan. But recent local campaigns in the energy sector require a change in policy.

Opposition to nuclear power has grown and several major projects — including the Minsk nuclear power and district heating station, and the Vitebsk, Chynyrn and Crimean nuclear power stations — have been cancelled or converted to conventional fuel. The Armenian nuclear power plant will be phased out by 1990, and further extension of the Ignalina power station in Lithuania has been cancelled. But these concessions to local opinion still

better than Soviet Union

remain decisions taken in Moscow. The governments of the individual republics still have no say in the development of nuclear power within their territories, and so far only the Lithuanian Supreme Soviet has decided to defy Moscow and withdraw its financial support from nuclear energy.

There is no clear alternative to the Soviet Union's commitment to nuclear power. Large hydroelectric dams are increasingly seen as uneconomic and harmful to the environment, except perhaps single dams on mountainous rivers. But in the mountains of Soviet Central Asia, the local population is alarmed about the seismic hazards of large dams. Proposed solutions such as solar energy in Moldavia or wind farms in Byelorussia seem somewhat utopian.

In the coming year, Soviet energy planners will have to address the difficult problem of restoring confidence in the All-Union energy programme.

Confidence in the space programme also needs to be restored. Space 'achievements' are increasingly viewed by the Soviet public as wasteful attempts to gain international prestige. The partial failure of the Phobos mission did not help to arouse enthusiasm for a manned mission to Mars some time between 2010 and 2020, and a new society to stimulate interest in space among young people has so far failed to take off. The socialist allies feel excluded by the new Soviet commitment to 'international' projects involving France and neutral and non-aligned nations.

Perestroika has set a challenge to the socialist bloc. Reactions range from those of Hungary, where the breakaway Democratic Trade Union of Academic Workers is now a recognized part of the political scene, and Poland, where the Academy of Sciences has just rejected a government report on the environment, to Cuba, which opposes the whole idea of *perestroika*, and East Germany, which claims that reform is unnecessary. In between come Bulgaria, where lisenkoism has only just been finally dispatched with a massive change of personnel in the Academy of Sciences, and Czechoslovakia, where calls for restructuring remain largely on paper.

An encouraging development in 1988 was the holding of international conferences by two new non-governmental organizations based in the socialist bloc — the Hungarian-based International Association of Physics Students and the Bulgarian-based 'Ecoforum'. Both meetings attracted wide support from East and West and were hailed by Soviet 'establishment' participants as a manifestation of the recent trend in 'informal' initiatives.

Vera Rich

Guarded optimism greets increased funds in Britain

London

BRITISH science enters 1989 on an optimistic note, even though 1988 saw no improvement in university finances nor in those of the research councils. The raised spirits in the research community stem from the money for science won from the Treasury by the Department of Education and Science. An extra £95 million is being added to the budget of the research councils for the year 1989–90, an increase of 16 per cent on 1988–89. Specific programmes will absorb much of that money, but, with the remainder, the research councils will be able to support some more research and fund a few more alpha-rated projects. Exactly how the money is to be spent will be announced soon.

Although the award is not extravagant, it is a real increase in science spending and reflects a change in the government's attitude towards science. That change was marked by a speech made to the Royal Society in 1989 by the prime minister, Margaret Thatcher, in which she expressed a commitment to the public funding of basic science.

But research dubbed 'near-market' suffered at the hands of the government in 1988, and the effects of this hit the research councils hard. A review of the work carried out by the Agricultural and Food Research Council (AFRC) identified research valued at £33 million a year as near-market. If the cuts are implemented in 1989, the council estimates that 750 jobs could be at risk.

The Natural Environment Research Council (NERC) was another victim of government policy regarding near-market research. In 1988, the council was forced to cut its support for university research by £2.6 million and to cancel the grants round for October. It announced 130 redundancies to take effect before April 1989, some of which will be compulsory. The council hopes that the worst is now over and that 1989 will be more stable. But the council's new chairman will be fighting to retain not just financial support but the council's identity as the year is likely to bring a merger with the AFRC.

The question of the structure of the research councils resurfaced in 1988 and two committees will report early in 1989 on what form support of the biological sciences should take. The House of Lords committee reviewing support for the biological sciences recommended a merger between the NERC and the AFRC in its interim report in 1988, but in its full report may recommend further evolution towards formation of a council for all the biological sciences which would incorporate some research of the Science

and Engineering Research Council as well as the Medical Research Council. A committee for the Advisory Board for the Research Councils will also give its view in the spring.

Biologists in the academic science community must be looking forward to 1989 with some trepidation. Following the reviews of the Earth sciences, chemistry and physics in universities, which were published in 1988, is the review of the biological sciences. A draft report is now ready; if, like the chemistry and physics reviews, it recommends a concentration of resources in departments of a minimum size, 1989 may witness the beginning of the end of the small science department.

Overseeing the restructuring will be the new Universities Funding Council (UFC), set up in 1988 to take over from the University Grants Committee. The new chairman of the council, Lord Chilver, is embroiled in controversy even before he takes up his post, declaring his view on access to higher education: courses should be available to those who value education highly enough to commit resources to it. His views will find favour in government, which made moves in this direction in 1988 with the publication of a policy document on student loans. Issues such as the repayment of loans and the involvement of commercial banks are still to be resolved, but the government aims to introduce the system in 1990. More protests from students — such as those resulting in clashes between students and police in the streets of London in 1988 — seem likely in 1989.

Dominating the work of the heads of universities in 1988 was the Education Reform Act. On the question of the abolition of tenure, the government emerged the victor — abolition will go ahead — but on the introduction of a system of 'contract funding', the universities' heads emerged victorious. They won a concession from the government that before any conditions are laid down on funds from the UFC, bodies representing the universities will be consulted. But the value of this victory was thrown into question when the Committee of Vice-Chancellors and Principals received from the Department of Education and Science a draft copy of the guidelines to govern the relationship between the government, the UFC and the universities, which described controls the government would have over universities' finances. While the government is encouraging universities to seek a broader base of funding, it seems determined to monitor those finances. Retaining their autonomy will still be high on the agenda of the universities in 1989.

Christine McGourty

History of 'biotechnology'

SIR—Although old uses of the word 'biotechnology' have been noticed¹, no satisfactory lineage has been traced. I should like to draw attention to a source that fulfills three strict criteria: a self-conscious claim to novelty, recognition from others of that novelty and noticeable use of the word.

In 1919, Kark Ereky, the Hungarian agricultural economist and, briefly, minister of food in the counter-revolutionary government, published an 84-page manifesto entitled: *Biotechnologie der Fleisch-, Fett- und Milcherzeugung in landwirtschaftlichen Großbetrieb*. Following the precedents of chemical technology and mechanical technology, Ereky coined his new word to cover the interaction of biology with technology, connoting all production by means of biological transformation:

Auf Grund des gleichen Gedankenganges weist der Verfasser alle die Arbeitsvorgänge, bei denen aus den Rohstoffen mit Unterstützung lebender Organismen Konsumartikel erzeugt werden, dem Gebiete der Biotechnologie zu. [original emphasis]²,

which may be translated as: "On the same grounds the author classifies as biotechnology all the lines of work by which products are produced from raw materials with the aid of living organisms".

Ereky's book was well received and its originality recognized. *Die Naturwissenschaften* described the author's goal as a new branch of knowledge that he calls *Biotechnologie*³.

Thereafter, the word biotechnology appeared in various contexts in the 1920s and early 1930s. Some applications clearly follow from Ereky's formulation. It appears immediately as a heading in *Zeitschrift für Technische Biologie*⁴.

Other uses should not, however, be attributed to Ereky. The similar word *Biotechnik* had come into use at about the same time, and although there is a difference in nuance, even in German the two words could be interchanged⁵. In English translation, the distinction is almost entirely lost. *Biotechnik* therefore spawned the alternative meaning of the English biotechnology. In one of the first English uses, *Nature* entitled a 1933 editorial "Biotechnology"⁶. Dealing with technology and the problem of maintaining the quality of the human species, its title seems to refer to the concept of *Biotechnik* presented by Goldscheid in his social biology of 1911⁷.

Since then the word biotechnology has been reformulated many times: in 1947 to describe ergonomics⁸, in 1962 as the title for a new journal dealing with biochemical and microbiological engineering⁹ and in 1979 to describe the potential of genetic manipulation¹⁰. The distinction between

different usages has often been overlooked because they do share an important connotation: a new biological approach to a great range of industries. This breadth was foreshadowed in the aftermath of the First World War by Ereky's formulation of the concept of biotechnology.

ROBERT BUD

*The Science Museum,
London SW7 2DD, UK*

1. Dixon, B. *New Scientist* **105**, 38 (1985).
2. Ereky, K. *Biotechnologie der Fleisch-, Fett- und Milcherzeugung im landwirtschaftlichen Großbetrieb* (Paul Parey, Berlin, 1919).
3. Pringsheim H. *Naturwissenschaften* **7**, 112 (1919).
4. Lindner P. Z. *Technische Biol.* **8**, 54–56 (1920).
5. Hase A. Z. *Technische Biol.* **8**, 23–45 (1920).
6. *Nature* **131**, 597–599 (1933).
7. Goldscheid R. *Höherentwicklung und Menschenökonomie: Grundlegen der Sozialbiologie* (Werner Klinkhardt, Leipzig, 1911).
8. Taylor, Craig L. & Boelter, L.M.K. *Science* **105**, 217–19 (1947).
9. Bud, Robert. *Chem. Brit.* **24**, 441–444, 466 (1988).
10. US Trademark, 1180658 filed 3 December 1979, expired 1 July 1988.

Primer availability

SIR—The possible applications of the polymerase chain reaction (PCR) (see ref. 1), which amplifies specific DNA sequences during successive cycles of DNA denaturation and extension of specific oligonucleotides (primers) by DNA polymerase, are enormous, both for basic research and for molecular diagnosis of inherited disease. In most laboratories involved in molecular diagnosis of genetic diseases, PCR will gradually replace routine DNA techniques of Southern blotting and hybridization with radioactive probes. A prerequisite to perform PCR, however, is the availability of the specific primers.

Recently, I requested primers for prenatal diagnosis of two common genetic diseases. On both occasions the answer was negative, although the papers describing the primer sequences and their possible applications had already been published. Costs of synthesis and distribution of the primers were probably the main factors hampering their release. If this becomes a general policy, laboratories will each have to face the problem of synthesizing the respective primers for all the different diseases they are studying.

Most companies charge \$1,000–\$2,000 to synthesize 10 micromoles of a 20–30 base pair oligonucleotide, which is enough to provide 100 laboratories with primer for 1,000 PCRs. With small fees (\$15–\$20) from each of the laboratories requesting the primers, costs for synthesis and distribution could easily be covered. It might be opportune, for logistic reasons, to centralize primer synthesis and distribution in an organization such as the American Type Culture Collection (ATCC), which has become a central depository and distributing organization

for DNA sequences.

The free release of DNA probes to laboratories throughout the world has contributed largely to the exponential increase in our understanding of genetic diseases. For the future, it is essential to guarantee the availability of the necessary PCR primers.

PATRICK J. WILLEMS

*Department of Medical Genetics,
University of Antwerp/UITA,
Universiteitsplein 1,
2610 Wilrijk, Belgium*

1. Saiki, R.K. *et al. Science* **239**, 487–491 (1988).

Soviet refusniks

SIR—We appreciate the increase in the number of exit visas from the Soviet Union granted to Jews, including some long-term refusniks who want to emigrate.

But despite the sizeable increase in the number of visas, a substantial number of long-term refusniks have no foreseeable hope of leaving. Several of them have been waiting for more than 15 years, and some have been told that they will not be allowed to leave until 1992 or 1996.

In most cases, the reason given is that of having access to secrecy. For instance, the Uspensky family has been refused because Igor Uspensky's mother, who is in her late seventies, worked in the Ministry of Agriculture until she retired 12 years ago. She has been told that the reclassification of her secrecy will take place after 1996 (when she will be 84 years old). This not only prevents her from leaving but also her son, daughter-in-law and grandsons. If they were not so tragic, one could put such classifications in the realms of the 'Theatre of the absurd'.

INGA FISCHER-HJALMARS

*International Federation of Scientists
for Soviet Refuseniks,
33 Seymour Place,
London W1H 6AT, UK*

Biblical criticism

SIR—The recent note (*Nature* **335**, 203; 1988) on the history of the survival of the Aleppo Codex describes the codex as the Hebrew Pentateuch, the first five books of the Old Testament. But the portion reproduced is of a passage from the Book of Judges (9:57 to 11:2). It includes the first part of the story of Jephthah — his exclusion from his paternal inheritance by his step-brethren because he was "born of another woman".

MYER H. SALAMAN

*23 Bisham Gardens,
London N6 6DJ, UK*

● The Aleppo Codex originally contained all the books which now constitute the Old Testament; about one-third of the text has been damaged or destroyed. Editor, *Nature*

Prospect for a new year

The temptation to guess what lies even immediately ahead must be resisted, as the prospects for fundamental physics show all too plainly.

PRETENDING to predict the future of discovery is not merely foolish, but a contradiction in terms. Discoveries are, of their nature, unexpected to some degree. What follows should therefore kindly be regarded not as a set of predictions but as a list of wishes. The exercise is worthwhile not for its content, but because it reflects the way that at any time there is a sense of expectation attached to some fields of research, and of dormancy to others. Perhaps a later Thomas Kuhn will think it worthwhile to compare these expectations with the realities that emerge.

Expectations are most vividly engendered by big machines, most of which are built to timetables advertised in advance. The year ahead should see the commissioning of two machines of great importance — the new electron-positron collider (LEP) at the European high-energy physics laboratory (CERN) at Geneva and the Hubble Telescope, which should at last be launched when the US shuttle is back in service and has room to spare. Curiously, neither machine has done as much to lift the spirits of those concerned as might have been expected.

The difficulty at CERN is that, this time, there is no obviously spectacular prize to win, such as the predicted discovery of the mesons that mediate the electro-weak force, and whose masses were predicted to be within the range of collisions between protons and anti-protons in CERN's super-synchrotron. Instead, as if discomfited that the top quark has not turned up in the same range of energy, and as uncertain as ever where the Higgs particle may lie, people are now hoping that 'new physics' will provide LEP's fun and excitement. Nobody belittles the importance of being able to measure the electro-weak bosons in plenty, but the hazard of new physics is the difficulty of telling the difference between the real thing and spurious data.

The disappointments of the Hubble Telescope are different: frustration and continuing uncertainty. The fact that an instrument which, if it works half as well as planned, is likely to do more to change our view of the Universe than all this century's optical instruments put together should still be sitting on the ground at Los Angeles after three years, waiting for a firm launch date, may help researchers appreciate their properly modest place in the wider scheme of things, but it speaks wonders for the endurance of the human

spirit in the face of frustration that the Space Telescope Science Institute and its principal investigators remain cheerful.

Some say the delay may have helped ensure success. The guide stars on which accurate pointing of the telescope will rely have been more accurately and comprehensively measured, while ground access has got rid of an optical misalignment together with a huge amount of unwanted dust. But the telescope has been frozen in design during a period when telescope designers have been more innovative than ever. There is also some concern that the modified shuttle may not be able to put the telescope into a safe orbit.

Either of these machines might help throw light on the great cosmological issues — but not in their first year. Appropriately, the Hubble Telescope should eventually provide a reliable distance scale for the Universe, and thus decide what Hubble's constant is, but everybody understands what painstaking work that must be. Even the issue of what quasars are will take time, though opinion does seem to be drifting towards the view that the nuclei of most kinds of active galaxies embody black holes; sadly, less attention is being given to questions such as the lifetime of the active phase and the related question of whether an active phase is an evolutionary stage of most, perhaps all, galaxies.

Expectations of the contributions likely to be made by particle physics to cosmological understanding are now less uniform than a few years ago, when the Big Bang seemed to many to be nature's practical demonstration that the particles of the known hierarchy (nobody calls them "fundamental" any more) are related to each other by the progressive breaking of a natural symmetry in a mass of expanding and thus cooling gas. Part of the trouble is psychological; no sooner had the particle theorists sketched in the details of the Big Bang in the first few seconds and minutes (with observationally important predictions of deuterium abundance and spatial uniformity in the early Universe) than, it seemed to cosmologists, they fell under the spell of string theory, which is a way of accounting for the differences between particles by means of dynamical variables in dimensions (most commonly, there are six of them) which are hidden from observation, or are "collapsed".

In principle, of course, particle physicists may justly say that the attractiveness

of string theory does not imply an abandonment of the exciting vision of how particles run down the hierarchy in the Big Bang, transforming from one species to another in the process, but merely a quest for a more natural explanation of the relationships between particles.

But they have nevertheless damaged the simple view that particle physics and cosmology would come together through the Big Bang in at least two ways. First, the mere suggestion that there may be unobservable dimensions is an unsettling reminder of other conundrums, the quantization of gravity, for example. (The question whether the hidden dimensions may embody the hidden variables for which Einstein fought is too little discussed.)

Second, if the dimensionality of space-time is to have a place on the agenda of discussion, should not attention also be paid to the arguments in the second half of Stephen Hawking's recent book *A Brief History of Time* that the Big Bang is an illusion brought about by occupancy of what is bound to seem a Universe in a corner of a larger structure?

It would be good if there were some progress in dealing with this muddle in the year ahead, because people's instincts are sound and the thinking will have to be done at some stage. On a less abstract but equally teasing plane, it does seem certain that 1989 will be the year of Berry's phase — the notion that for any mechanical system represented by a complex field (such as a common-or-garden electronic wave function), whose phase is supposed to have no physical significance, each characteristic of the system can be used to define a phase in such a way that supposedly identical states have different phases, which may be measured as, or inferred to be, different. Abrahamov and Bohm, independently of M.J. Berry, have shown that the supposedly insignificant complex phase of Maxwell's electromagnetic potential is measurable. Is this another way of importing hidden variables? People are now so busily generalizing Berry's argument that we should know what it means before the year is out. The Nobel physics committee would be well advised to plump for this trio before it is denied the topic by an embarrassment of choice. After all, the looked-for explanation for high-temperature superconductivity cannot now be found before 1990.

John Maddox

Cell biology

Reception and transmission

Asser S. Andersen

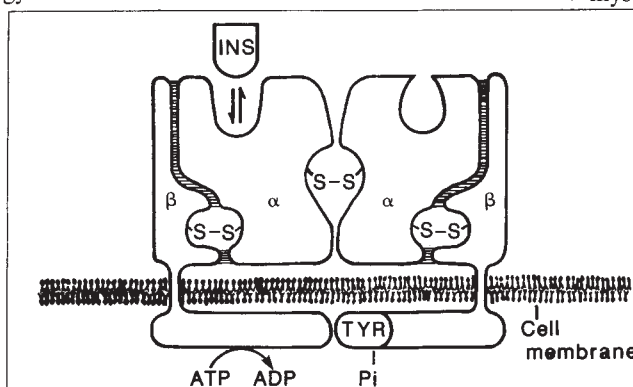
THE activity of cell-surface receptors is in many cases controlled by their association state in the membrane, and the association state is in turn regulated by the attachment of the ligand. Rutter and his colleagues in San Francisco now show¹ that this ligand-dependent association can be mimicked in free solution by the extracellular domain of the insulin receptor. The precise relationship between receptor association and the biochemical events that the receptor controls in the cell is less well-defined, but here too there are some new pointers.

Cell biology and endocrinology have both come a long way since the first evidence that insulin can regulate the activities of a cell without actually going inside was met with incredulity. Signal transduction has now become almost a discipline of its own, and the insulin-insulin receptor system remains the prime example. The way it works is as follows. The insulin receptor, embedded in the cell membrane, is a heterotetramer of four subunits ($\beta\alpha\beta\alpha$) held together by disulphide bonds (see figure). Insulin binds tightly to the extracellular α -subunits, causing the activation of a tyrosine kinase activity which resides in the intracellular domain of the β -chains. The β -subunits then phosphorylate each other, leading to a rise in kinase activity. A still largely uncharacterized regulatory cascade is set in train, involving the phosphorylation of several cytoplasmic proteins. A review of these processes published some years ago was entitled "Insulin binds, and then something happens": although the frontiers of ignorance have been pushed back since then, there are areas that are still shrouded in fog.

To cell biologists, the question of greatest interest must still surely be how the binding of the insulin molecule activates the receptor, and here there has recently been great progress, much of it based on the introduction of alien insulin receptors into mammalian cells in culture. The α - and β -chains are synthesized *in vivo* on a single messenger RNA and then proteolytically separated and inserted into the cell membrane.

Johnson *et al.* have transfected cells of a mammalian (CHO) cell line with complementary DNA encoding either the whole $\alpha\beta$ ensemble or truncated versions lacking most of the β -chain, so as to eliminate both the transmembrane and kinase

domains. These secreted 'ectodomains' emerge with the correct quaternary structure (two disulphide-linked α -chains, each with its vestigial β -chain), and properly finished in terms of glycosylation. They bind insulin with fully developed affinity (which is not surprising, for the insulin-binding domain is intact). The striking observation of Johnson *et al.*¹ is that when insulin is added to a solution of purified ectodomains, these entities associate into small clusters or, depending on the concentration, long aggregates, which branch or loop back on themselves. The authors infer from the similar effect of



The subunit structure of the insulin receptor. (Adapted from ref. 8.)

epidermal growth factor (EGF) that the aggregation is a consequence of mere occupancy of a ligand site, rather than a specific effect of insulin. (On the other hand, although the EGF and insulin receptors are structurally related, and EGF does perturb biochemical responses to insulin in some cell types², it has never been found actually to compete with insulin on the insulin receptor.)

Other groups have also tinkered with insulin receptors. In particular, Maegawa *et al.*³ have introduced a truncated recombinant version of the human insulin receptor, devoid of kinase activity, into rat fibroblasts which have the normal (or even an increased) complement of functional receptors. It turns out that these alien and silent receptors do not interfere at all with the action of the endogenous receptors, up to the point of phosphorylation of the β -chain and stimulation of its kinase activity. But thereafter, things go wrong: the endogenous receptor fails to phosphorylate its cytoplasmic substrates, and the transfected cells show none of the normal physiological responses to insulin binding. An interesting (but by no means the only) explanation could be that receptor aggregates are needed to generate biological activity, and complexes of the active with the inactive sort are not

functional. The CHO cells that make human insulin receptors in addition to their own¹ have yielded another morsel of evidence for the importance of receptor aggregation⁴: when there are more receptors, fewer insulin molecules are needed to generate a given response. This implies that the closer the receptors are in the plasma membrane the better they amplify the binding signal.

Using the more homespun methods of traditional biochemistry, the groups of Pessin⁵ and of Kahn⁶ have cast new light on the early steps of insulin action. Sweet *et al.*⁵ find that mild reduction causes the heterotetramer to dissociate into $\alpha\beta$ -dimers. When incubated with insulin under oxidizing conditions, these revert to the tetramer, showing again that extracellular ligand regulates association state.

Shoelson *et al.*⁶ re-examine an old mystery — that trypsin can mimic the effects of insulin in various cells. Their observations suggest that a proteolytically induced relaxation of the α -subunit releases the steric constraints which normally keep the β -subunits from phosphorylating each other. Such a model is consistent with the properties of another engineered fragment: a truncated insulin receptor, lacking the insulin-binding domain, confers the complete gamut of insulin responses on the cell in which it is expressed, without the need for insulin⁷. So it seems that the function of the α -subunit is to keep the β -subunits dormant when there is no insulin to press the switch.

All this is progress indeed, but questions remain to be answered. For instance, when receptors are isolated from most cell types they appear to be exclusively in the $\alpha_2\beta_2$ form and phosphorylation, which is presumed to take place within the tetramer on binding of insulin⁵, requires no further aggregation. What then are we to make of the aggregates of Johnson *et al.*? Is it possible that EGF can modulate the cell's response to insulin by altering the association of the insulin receptors? Does phosphorylation or some other kind of interplay between heterotetramers control a later event in the insulin response? □

1. Johnson, J.D., Wong, M.L. & Rutter, W.J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7515-7520 (1988).
2. Hollenberg, M.D. & Cuatrecasas, P. *J. biol. Chem.* **250**, 3845-3853 (1975).
3. Maegawa, H. *et al. J. biol. Chem.* **265**, 12629-12637 (1988).
4. Ebina, Y. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 8014-8018 (1985).
5. Sweet, L.J. *et al. J. biol. Chem.* **262**, 16730-16738 (1988).
6. Shoelson, S.E., White, M.F. & Kahn, C.R. *J. biol. Chem.* **263**, 4852-4860 (1988).
7. Ellis, L. *et al. Molec. Endocrin.* **1**, 15-24 (1987).
8. Pilch, P.F. *et al. in Insulin Action and Diabetes* (eds Goven, H.J. *et al.*) **74** (Raven, New York, 1988).

Asser S. Andersen is in the MRC Cell Biophysics Unit, 26 Drury Lane, London WC2B 5RL, UK.

Computation

Highly distributed processing

Ian Stewart

WHAT is the best way to obtain raw computational power? A massive central supercomputer or a network of smaller machines? Both have their passionate devotees. A recent calculation in number theory has given a new meaning to the phrase 'distributed processing': it illustrates the possible advantages of a network for problems that can be broken into independent pieces. A 100-digit number, carefully screened for maximum difficulty, has been resolved into prime factors by the communal efforts of a dozen mathematicians in three countries using 400 separate computers simultaneously and communicating by electronic mail. The marathon calculation widely reported in the popular press (see *Nature* 335, 658; 1988) was organized by Mark S. Manasse of the Digital Equipment Corporation's research centre at Palo Alto and Arjen K. Lenstra of the University of Chicago. It took less than a month: a supercomputer working non-stop would have taken at least a week and cost a great deal more.

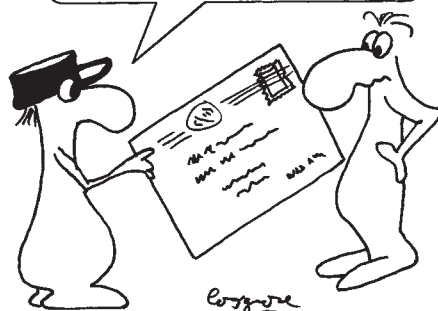
The factorization of large numbers into primes used to be of purely theoretical interest. Today, however, the problem has important applications in cryptography and cryptanalysis. A well known method of encoding sensitive information, the RSA (Rivest-Shamir-Adelman) system, relies for its security on the conjectured difficulty of resolving an arbitrary number into its constituent primes. Any improvement in factorization methods is thus automatically of interest in a much wider sphere. The communal effort under discussion is in a sense most noteworthy for its *lack* of novelty. The actual method employed, known as the quadratic sieve, was invented by Carl Pomerance in 1982 (*Computational Methods in Number Theory Part I* (eds Lenstra, H. W. Jr & Tijdeman, R.) Mathematisch Centrum Tract 154, Amsterdam; 1982), and was rapidly accepted as one of the fastest methods available.

The first thing to understand about prime factorization is that the most obvious method, trial division of the number N by each possible prime in turn up to the square root, is hopelessly inefficient for numbers with more than just a few digits. The number of primes less than $\sqrt{N}$ is approximately $2\sqrt{N}/\log N$, so the number of trials needed for a 100-digit number is roughly $2 \times 10^{50}/(50 \log 10) = 1.7 \times 10^{48}$. At a million trials per second the computation would take about 5×10^{34} years, or at least 10^{27} times the age of the Universe. Variations on this technique, such as strategies for searching specific ranges of

primes, like starting at $\sqrt{N}$ and working backwards, suffer from the same defect.

The solution is to apply more subtle methods, based on results in number theory. Suppose we wish to find the prime factors of a number N . If we can find just one of them, we can divide it out and repeat the process on the much smaller number that results. So the real problem is to find *some* prime factor. One approach is to look for two perfect squares whose

The final step in your multimillion-dollar programme to solve Fermat's Last Theorem has arrived, Professor Higgins. Special Delivery. Held up two years by a postal strike in Zimbodia.



difference is a multiple of N . Let x^2 and y^2 be these squares. Because $x^2 - y^2 = (x - y)(x + y)$, any prime factor of N must divide either $x - y$ or $x + y$. Suppose for the sake of argument that it divides $x + y$; then it also divides the highest common factor of $x + y$ and N . There is a very efficient method, known as the euclidean algorithm, for computing this highest common factor, which in general is much smaller than N . Because it is smaller, we can find its prime factors relatively easily; but these are also prime factors of N . The strategy is thus very much one of 'divide and conquer'.

It's a neat trick — but how do we find x and y ? Michael Morrison and John Brillhart (*Mathematics of Computation* 29, 183–205; 1975) realized that suitable candidates for x and y could be found by using the continued fraction for $\sqrt{N}$, a sequence of fractions p/q that approximate $\sqrt{N}$ ever more closely. Their proposal was to run along this sequence of fractions waiting for the denominator q to become a perfect square. When this happens, suitable x and y can easily be found. But there is a much better way. Instead of waiting for a square denominator to turn up, we can try to create one. Some upper limit is chosen and each denominator is checked to see whether all of its prime factors lie below this limit. If so, its factorization is stored as a sequence of zeros and ones: 0 if the

corresponding prime factor occurs to an even power, 1 if it occurs to an odd power. In this way a huge matrix is built up, with these sequences as its rows.

Some subset of all these denominators, multiplied together, will form a square — provided that the corresponding sequences have, between them, an even number of ones in each column. This can be checked quickly by standard techniques of linear algebra, known as gaussian elimination. Having found a combination of denominators whose product is a perfect square, numbers x and y such that N divides $x^2 - y^2$ are again easy to find. This is the basic idea: the full-blooded quadratic sieve incorporates several further refinements, including strategies to cut short the computation if it is getting nowhere. (For detailed description see *Prime Numbers and Computer Methods for Factorization* by Hans Riesel; Birkhäuser, Boston, 1985).

Manasse and Lenstra's collaborators were from the United States, the Netherlands and Australia. The matrix required to apply the quadratic sieve turned out to have 50,000 rows and columns, a total of 250 million entries. The computation of the matrix was split into 400 pieces, the results were assembled by electronic mail at Palo Alto, and the final gaussian elimination led to the discovery that the number had two prime factors, with 40 and 61 digits. There are so many computers now that much of their time is effectively free, especially for low-priority jobs running 'in background', and the tale bears dramatic witness to the capabilities of a network of small machines running in parallel.

The quadratic sieve is a highly refined and extremely clever technique, but has by now become fairly standard. So in a sense the method is an example of brawn rather than brains. The particular style of brawn, large-scale international cooperation, is seldom applied in such an organized manner. On the other hand, the whole of mathematics is a large-scale international cooperation, whose customary methods of communication are often a good deal slower than electronic mail.

By pushing the present approach to its limits, factorizations of 110- or perhaps 120-digit numbers might be achieved. To factorize an arbitrary 150-digit number, say, still looks out of the question, and the only hope is a completely new idea. This is not inconceivable; for example Hendrik Lenstra (of Amsterdam University) recently suggested a new approach involving so-called elliptic curves (see my News and Views article in *Nature* 325, 199; 1987), which can sometimes be extremely efficient. But no known method works efficiently for the awkward cases as well as the nice ones. In the world of prime factors, nobody yet knows whether brain will ultimately prove superior to brawn. □

Ian Stewart is in the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

RNA processing

Splicing a spliceosomal RNA

David A. Brow and Christine Guthrie

Two fundamental questions that captivate the interest of those who study RNA processing are how interrupted genes acquired their introns and how the removal of introns from RNA is catalysed. Tani and Ohshima report on page 87 of this issue¹ a remarkable finding that may bear on both questions: they identify an intron within a gene encoding a spliceosomal RNA. The spliceosome is the structure responsible for removal of introns from nuclear pre-messenger RNA (Fig. 1). It is assembled from the pre-mRNA, four small nuclear ribonucleoprotein complexes (snRNPs) called U1, U2, U5 and U4/6, containing five small nuclear RNAs (snRNAs), and an as yet unknown number of additional proteins (see ref. 2 and the News and Views article³ by M. R. Green). Because the intron discovered by Tani and Ohshima has the distinctive structure of nuclear pre-mRNA introns, it seems in this case that an snRNA is spliced by the very apparatus of which it is an important component.

The intron interrupts the single-copy gene encoding U6 snRNA in the fission yeast *Schizosaccharomyces pombe*, and contains perfect matches to the consensus sequences that have been identified at the splice sites and at the site of branch formation in nuclear pre-mRNA. Tani and Ohshima point out¹ that this strict adherence to the consensus motif, in an organism where these sequences are often rather divergent, suggests the U6

intron is under selective pressure for efficient removal. Indeed, Tani and Ohshima observe only fully spliced U6 snRNA *in vivo* under normal growth conditions. Although the participation of U6 in the removal of the U6 intron can as yet only be inferred, a consideration of the possible origin of this intron has interesting implications for the general role of U6 snRNA in splicing.

It is the snRNAs, rather than the proteins, that are believed by many to be the main catalytic elements of the spliceosome. This notion follows from the discovery of another class of RNAs, the group II self-splicing pre-mRNAs, in which excision of the intron and ligation of the remaining exons occurs via a pathway that is mechanistically very similar to that of nuclear pre-mRNAs, but does not require proteins^{4,5}. Indeed, it has been suggested that the spliceosome evolved from such autocatalytic RNAs⁶. Thus, in the study of spliceosome-mediated intron removal, much effort is being directed towards attributing specific catalytic functions to the snRNAs.

U6 snRNA is by far the most highly conserved spliceosomal RNA⁷, and so is expected to have a fundamental role in splicing. Yet, unlike the U1, U2 (and probably) U5 snRNA-containing snRNPs, which recognize the conserved intron elements (Fig. 1), the function of the snRNP containing U4 and U6 snRNAs is unknown. U4 and U6 are strongly associated with

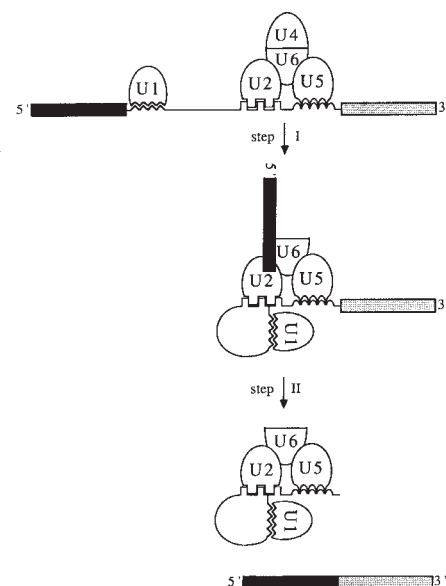


Fig. 1 Nuclear pre-mRNA splicing pathway. The 5' and 3' exons are represented as dark and light boxes, respectively, and the intron as a thin line. Top, fully assembled spliceosome, with U1, U2 and U5 snRNA-containing snRNPs interacting with conserved sequences at the 5'-splice site, branch point and 3'-splice site, respectively. Middle, during or before the first nucleolytic step of splicing, U4 dissociates from the spliceosome, which now contains free 5' exon and the branched 'lariat' intermediate. Bottom, step II produces spliced mRNA and free intron lariat, still bound by snRNPs. (Contacts between snRNPs do not imply proven associations.)

one another by base-pair interactions that form two contiguous stems (Fig. 2, stems I and II). Curiously, these stems lie within the most highly conserved segment of U6, consisting of 49 nucleotides with at least 80 per cent identity between yeasts and mammals. Normally, paired bases exhibit phylogenetic covariation, changing their identity but maintaining complementarity (for example, A-U changes to G-C). This paradox could be resolved by the surprising observation that the stable U4/6 interaction is disrupted during splicing *in vitro*, as judged by the loss of U4 from the spliceosome before or during the first nucleolytic step⁸⁻¹⁰ (Fig. 1). Thus, the conservation of U6 within the stems could reflect an additional role for this region when it is unpaired. Indeed, the timing of U4 dissociation with respect to the generation of splicing products suggests that this highly conserved region of U6 may be involved in catalysis. U4 could thus be viewed as an anti-sense RNA that negatively regulates U6 activity.

The intron identified by Tani and Ohshima is located precisely in the middle of the highly conserved region of U6, bifurcating interaction stem I (Fig. 2). Its presence in a U6 gene is particularly surprising as this is the first example of a pre-mRNA-like intron in an RNA polymerase III transcript^{11,12}. The fact that half a dozen

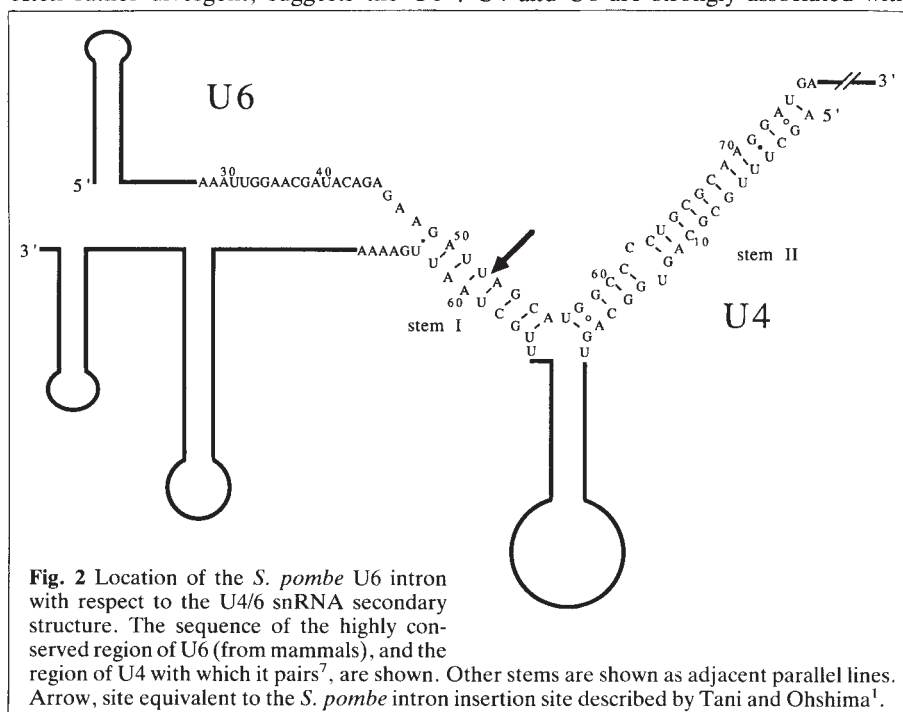


Fig. 2 Location of the *S. pombe* U6 intron with respect to the U4/6 snRNA secondary structure. The sequence of the highly conserved region of U6 (from mammals), and the region of U4 with which it pairs⁷, are shown. Other stems are shown as adjacent parallel lines. Arrow, site equivalent to the *S. pombe* intron insertion site described by Tani and Ohshima¹.

other sequenced U6 genes (including two from the budding yeasts) contain no intron, together with the obvious 'chicken-or-egg' problem presented by an intron in a spliceosomal RNA, suggest that the intron was not present in the progenitor U6 gene, but may have been acquired more recently in evolution. Given that the U6 intron does not appear to be a transposable element¹, how did it arise?

An intriguing possibility is that the *S. pombe* U6 intron arose from a mishap during pre-mRNA splicing. Perhaps, during its excision from pre-mRNA, an intron inserted into the U6 RNA molecule present in the spliceosome. The resulting product was presumably reverse-transcribed and its complementary DNA then replaced the chromosomal U6 gene. If this scheme is correct, the site of intron insertion would probably reflect the proximity of that part of U6 to the catalytic centre of the spliceosome. The picture of U6 that emerges is of a central, possibly catalytic site flanked by conserved sequences positioned to interact, directly or through other snRNPs, with the pre-mRNA near the 5' and 3' splice sites.

Insertion of an excised intron into RNA is not unprecedented: it has recently been found that autocatalysed splicing of a group I pre-ribosomal RNA (which

proceeds by a mechanism somewhat different from that used by the group II and nuclear pre-mRNA introns²) can run backwards, with the intron reinserting at the site of its excision (S. Woodson and T. Cech, personal communication). Although the origin of the intron discovered by Tani and Ohshima is a matter for conjecture, a significant step in advancing these ideas from the realm of speculation may come from further sequence analysis. Identification of an intron at the same site in the U6 gene of an organism unrelated to *S. pombe* would support the proposed mechanism of insertion and its implications for U6 function. □

1. Tani, T. & Ohshima, Y. *Nature* **337**, 87–90 (1989).
2. Guthrie, C. & Patterson, B. A. *Rev. Genet.* **22**, 387–419 (1988).
3. Green, M. R. *Nature* **336**, 716–718 (1988).
4. Peebles, C. L. *et al.* *Cell* **44**, 213–223 (1986).
5. van der Veen, R. *et al.* *Cell* **44**, 225–234 (1986).
6. Cech, T. *Cell* **44**, 207–210 (1986).
7. Brow, D. A. & Guthrie, C. *Nature* **334**, 213–218 (1988).
8. Pikielny, C. W. *et al.* *Nature* **324**, 341–345 (1986).
9. Cheng, S.-C. & Abelson, J. *Genes Dev.* **1**, 1014–1027 (1987).
10. Lamond, A. I. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 411–415 (1988).
11. Kunkel, G. R. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 8575–8579 (1986).
12. Reddy, R. *et al.* *J. biol. Chem.* **262**, 75–81 (1987).

David A. Brow and Christine Guthrie are in the Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA.

Submillimetre astronomy

Different views of the Universe

Glenn J. White

ONE of the few remaining unexplored regions of the electromagnetic spectrum accessible to ground-based astronomical observations lies at wavelengths between 1 and 0.3 mm. Using the new generation of large submillimetre-wave telescopes which have recently been completed in Hawaii, Europe and Chile, astronomers have been actively studying a wide range of astrophysical problems related to the Solar System, the physics of star formation, the chemical and dynamical evolution of the interstellar medium, the structure and evolution of galaxies, and the nature and energetics of active galactic nuclei and quasars. This spectral region is well suited to the study of gas and dust in the interstellar medium, particularly in its ability to sample regions which are highly opaque at optical wavelengths. The first results from these new telescopes, especially those relating to the interstellar medium, were discussed at two recent meetings*.

An accurate determination of the mass of material in the interstellar medium is crucial to many astrophysical studies, including those of the structure and

dynamics of molecular clouds and their subsequent formation of stars. Most of this mass is in the form of hydrogen, which remains very difficult to observe from the surface of the Earth. Consequently it is necessary to infer its presence from observations of other molecules such as carbon monoxide, which undergo frequent collisions with it. These collisions excite rotational (*J*) transitions in the molecular species, which are observable throughout the millimetre and submillimetre wavelength part of the spectrum. As part of the continuing controversy about the $N(\text{H}_2)/W(\text{CO})$ ratio (column density of H_2 over integrated intensity in the $\text{CO } J=1-0$ transition) in interstellar gas in our Galaxy and external galaxies, new observations indicate that the value of this ratio is about six times larger in the Large Magellanic Cloud than in our own Galaxy (T. Dame, Harvard-Smithsonian Center for Astrophysics), and up to 25 times larger in the Small Magellanic Cloud (M. Rubio, Harvard-Smithsonian Center). To explain these discrepancies, it was suggested that scattering of CO radiation in the low-density envelopes of clouds may in part account for the discrepancies between $N(\text{H}_2)$ determinations from

CO , ^{13}CO and C^{18}O in galactic clouds (M. Guélin, Institut Radio Astronomie Millimétrique).

In addition to the hydrogen gas, about 70 other molecular species have been detected in the interstellar medium as minor constituents. These are subject to a complex range of chemical interactions, which also influence the thermal balance of the gas and ultimately the structure and stability of molecular clouds. With the extremely high angular and spectral resolution now available in the millimetre and submillimetre wavelength region, detailed studies are possible which trace variations of both chemistry and excitation throughout molecular clouds. New evidence shows that ultraviolet radiation can penetrate interstellar clouds to a much greater extent than previously realized, because of the clumpy nature of material in molecular clouds. This can lead to an increase in the rate of photodissociation of the molecules, and modification of the local chemical equilibrium. These results, based on submillimetre wavelength observations of C I and C II (neutral and singly ionized carbon) emphasize the increasing importance of high-frequency observations (J. Stutzki, Max Planck Institut für Astrophysik, Garching).

One important consequence of this higher radiation field is to speed up the chemical processing and evolution of molecular clouds (W. Langer, Princeton University). The models, however, are still unable to explain all the recent observational estimates of the abundances of various molecules. Consequently, the importance of time-dependent gas-phase chemistry (E. Herbst, University of Cologne), and grain-surface reactions, particularly for hydrogenated amorphous carbon (W. Duley, York University) were discussed to account for the long-chain hydrocarbon molecules observed in regions like the Taurus molecular cloud.

After the gravitational collapse of a molecular cloud, stars may ultimately be born out of the gas, with some of the residual debris going on to form 'planetary' systems. Although it is perhaps remarkable that ground-based observations of rather 'average' stars in the Galaxy can still hold many new surprises, recent observations have confirmed the presence of dust shells around stars such as Vega, α PsA and β Pic (E. Becklin, University of Hawaii). The submillimetre wavelength fluxes, however, fall well below those predicted from an extrapolation of the Infrared Astronomical Satellite (IRAS) measurements, suggesting that the particles are mostly less than 100 μm diameter. The most likely explanation is that they have been generated since the formation of the star, probably as a consequence of cometary or asteroid collisions.

For molecular clouds to collapse to form stars, a delicate balance between

* *The Physics and Chemistry of Interstellar Molecular Clouds*, Zermatt, 22–25 September 1988 and *Millimetre and Submillimetre Astronomy*, Kona, Hawaii, 3–6 October 1988.

their internal pressure and the opposing force of gravity regulates the rate at which stars will be formed. Until recently it has been extremely difficult to estimate the strength of magnetic fields inside dense regions, although this has now become possible at millimetre and submillimetre wavelengths, in particular by measuring the amount of polarized radiation emitted from the interstellar dust. Measurements of the magnetic fields in several molecular clouds suggest that the magnetic contribution to the internal motions and energy of many of these clouds is crucial for cloud dynamics, cloud evolution and star formation (P. Myers, Harvard-Smithsonian Center; R. Hildebrand, University of Chicago). In addition it has been suggested that hydromagnetic waves may be responsible for inducing density fluctuations throughout the interstellar medium (R. Pudritz, McMaster University), and examples of large-scale filamentary and ordered structures can now be seen in extensive maps of regions such as the Orion molecular complex (J. Bally, AT&T Bell Laboratories). The size scales of these large structures, which can extend over many tens of square degrees on the sky, emphasize the unique role which can be played by small telescopes, whose large beams make it possible to map extensive areas of the sky rapidly (P. Thaddeus, Harvard-Smithsonian Center).

The Galaxy is a large place, and it remains difficult to build up a detailed view of the large-scale processes occurring during stellar formation. However, the physics of star-formation across a whole galaxy can be uniquely viewed in nearby

face-on spiral galaxies. Using a millimetre wavelength interferometer, new observations of M51 have been made which reveal the presence of giant molecular complexes having sizes of around 400 parsecs and masses greater than 10^7 solar masses. The velocity structure of the molecular material shows systematic gradients across individual galactic spiral arms, where the star formation efficiency is also found to be strongly enhanced. This suggests that density waves are responsible for triggering star formation on a galactic-wide scale (S. N. Vogel, Rensselaer Polytechnic Institute; see Vogel, S.N., Kulkarni, S.R. & Scoville, N.Z. *Nature* 334, 402-406; 1988).

The millimetre and submillimetre wavelength regions are important to the study of objects such as ultraluminous galaxies. These objects, first detected by IRAS, may in some cases go on to form quasars. Using a synthesis of recent results and new observations in the CO line, it was suggested that the collision, or merger, of two spiral galaxies, can result in removal of angular momentum and allow an infall of gas towards their centres. This rapid accretion leads to a star burst in the nucleus, and then the aggregation of gas and stars into a massive central stellar cluster. The subsequent nucleation of this core, and a continued infall of material from the outer parts of the merger may lead on to the formation of an active galactic nucleus or quasar (N. Z. Scoville, California Institute of Technology). □

Glenn J. White is in the Department of Physics, Queen Mary College, University of London, Mile End Road, London E1 4NS, UK.

Low-temperature physics

Viscosity of ^3He - ^4He solutions

P.V.E. McClintock

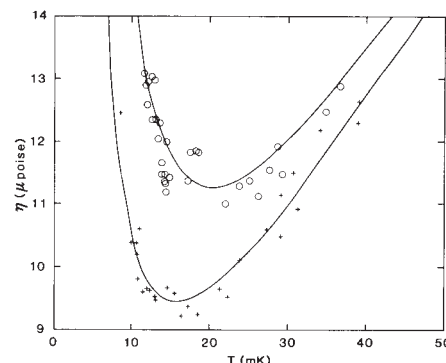
LIQUID helium thickens slightly when subjected to a strong magnetic field at very low temperatures. The latest measurements of this intriguing effect, which occurs only in liquid ^3He - ^4He isotopic mixtures or in pure liquid ^3He , are reported in a recent issue of *Physical Review Letters* (61, 1619-1622; 1988) by J.R. Owers-Bradley and colleagues at Nottingham University.

The new experiments were performed on a dilute solution containing 0.1 per cent of the rare helium isotope ^3He dissolved in the common isotope ^4He . The properties of liquid ^3He and liquid ^4He at low temperatures are markedly different, and a solution of this kind is a rather strange material. Liquid ^4He , a so-called boson system in which there is no restriction on the number of particles that may occupy any given energy state, becomes superfluid below a temperature of 2.17 K: it loses all

its viscosity and objects can move through it without experiencing the drag found in conventional liquids.

Liquid ^3He , on the other hand, is a fermion system in which there is a rigorous restriction that forbids more than one particle from simultaneously occupying any given energy state. (It too becomes superfluid, but must be a thousand times colder before the transition occurs.) When dissolved in ^4He , ^3He atoms move effortlessly through the superfluid. In fact, the superfluid ^4He just constitutes a kind of ethereal background, often termed a mechanical vacuum. The ^3He atoms in it form an almost ideal gas and, for most purposes, the presence of the superfluid background can be ignored.

Because of the restriction on the occupation of energy states, it is characteristic of fermion systems that they possess a great deal of energy, even at 0 K. The



Measurements of the viscosity η of a 0.1-per-cent ^3He - ^4He solution as a function of temperature T in zero magnetic field (+) and in a magnetic field of 10 tesla (○). The lower solid curve represents a fit of theory to the data with the strength of the interaction between the ^3He atoms as an adjustable parameter; the upper curve is a prediction by the same theory without any further adjustable parameters. (From Owers-Bradley *et al.*)

system tries to get into the lowest allowed energy configuration, which means that every permitted state is occupied from zero up to a maximum energy known as the Fermi energy, E_F ; all states above E_F are empty. The magnitude of E_F depends on the density of fermions and, in the case of a ^3He - ^4He solution, can conveniently be adjusted simply by changing the concentration of ^3He .

The ^3He nuclei carry tiny magnetic moments which, according to quantum mechanics, align either with or against an applied magnetic field, the two alignments having different energies. When a magnetic field is first applied to the cold liquid, half of the ^3He atoms suddenly have their energy increased relative to the other half so that, in effect, each group momentarily has a different Fermi energy. Relaxation processes then occur in which some of the higher-energy atoms make transitions to the lower-energy aligned states, until the value of E_F is again common to both groups. The liquid becomes polarized, there being more atoms aligned with the field than against it. At sufficiently low temperature, the degree of polarization achievable by this 'brute force' method depends on the relative magnitudes of the magnetic and Fermi energies. By using a dilute ^3He - ^4He solution, making E_F relatively small, Owers-Bradley *et al.* obtain a much higher degree of polarization in a given magnetic field than was the case in the earlier experiments of, for example, G.A. Vermeulen *et al.* (*Phys. Rev. Lett.* 60, 2315-2318; 1988) using pure liquid ^3He .

Because the interactions between pairs of ^3He atoms depend on whether their nuclear magnetic moments are parallel or antiparallel to each other, the properties of a polarized ^3He - ^4He solution are expected to differ markedly from those of an ordinary unpolarized one. The interaction is stronger between pairs of atoms whose moments are antiparallel, and these

obviously become fewer as the solution polarizes. The mean free path of the atoms consequently increases with polarization, generating corresponding changes in the transport coefficients, including an increase in viscosity. From the measured magnitude of the change in viscosity it should be possible to derive parameters of the interaction between the ^3He atoms.

Following a technique described by D.S. Greywall and M.A. Paalanen (*Phys. Rev. Lett.* **46**, 1292–1295; 1981), Owers-Bradley *et al.* used measurements of second sound — a density wave in the gas of ^3He atoms — to deduce the viscosity. They investigated standing waves of second sound in a resonant cavity and, after making appropriate allowance for other dissipative processes, determined the viscosity of the liquid from the measured widths of the resonances. The results (see figure) show clearly that the viscosity increases on application of the field, the increase becoming substantially larger as the temperature is reduced. The results should be more accurate and reliable than the earlier ones of Greywall and Paalanen who worked with a resonator which produced a distorted line shape owing to a hole in the cavity wall needed to assist cooling of the sample inside. In the new resonator, the inner wall is made of sintered silver, whose large surface area allows the sample to be cooled directly without any need for a hole to the exterior refrigeration arrangements.

An interesting feature of the results is that, unlike in an earlier experiment by D.A. Ritchie *et al.* (*Phys. Rev. Lett.* **59**, 465–468; 1987), there seems to be no evidence of 'slip' at the walls: rather, there was a boundary layer of effectively stationary ^3He at the walls, as would be expected with an ordinary viscous fluid. The authors suggest the difference may arise because Ritchie *et al.* used a smooth-walled chamber, quite different at the microscopic level from the rough, porous, sintered silver walls of the present experiment.

The lower curve of the figure represents a fit of a theory, based on ideas originally developed by E.P. Bashkin and A.E. Meyerovich (*Adv. Phys.* **30**, 1–92; 1981), to the zero-field viscosity measurements, using the strength of the interaction between the ^3He atoms as an adjustable parameter. The agreement is excellent and in itself suggests that the deduced form of the interaction is probably correct. The real test, however, is seen in the upper curve which is a prediction of the behaviour at 10 tesla without recourse to any additional adjustable parameters. Again, the agreement is excellent, apparently confirming the validity both of the theory and of the interaction deduced by fitting it to the zero-field data. □

P.V.E. McClintock is a reader in physics at the University of Lancaster, Lancaster LA1 4YA, UK.

Sickle cell anaemia

A simple disease with no cure

L. Luzzatto and P. Goodfellow

MOLECULAR biology has explained the basic defect in sickle cell anaemia and has provided the tools for prenatal diagnosis, but the treatment of afflicted patients undergoing crisis has not correspondingly improved. For the pessimist, this fact poses the general question of the efficacy of molecular approaches for treatment of genetic disease; and this echoes previous debates about the reductionist approach of molecular biology. Be that as it may, a positive step was taken by the Wellcome Trust in the sponsoring of a Sickle Cell Workshop, held in London on 4 November 1988, to consider new approaches to the treatment of sickle cell anaemia.

No genetic disease could be simpler than sickle cell anaemia. A single base pair in the β -globin gene is mutated from adenine to thymine producing a corresponding change in the sixth amino acid of the protein from glutamic acid to valine. Under conditions of low oxygen tension, the abnormal deoxyhaemoglobin polymerizes and eventually leads to a change in red blood cell shape termed sickling. When distorted red blood cells occlude small blood vessels, this can result in 'crisis', a term that dramatically describes a wide spectrum of events, from relatively brief painful episodes to massive thrombosis leading to infarction in bone, lung or brain. The organ damage is potentially crippling and the crisis may be life-threatening.

Following the theme of "know your enemy", G. Serjeant (MRC Laboratory, Jamaica) described the natural history of sickle cell anaemia. The trigger for crisis is not clear, but its onset is often associated with a drop in environmental temperature (Redwood, A.M. *et al. Br. med. J.* **i**, 66; 1976). Exposure to cold may divert blood flow from the bone marrow, causing local necrosis and 'bone' pain. Exploration of this hypothesis will require quantitation of the environmental factors triggering crisis and measurements of blood flow to the bones of patients suffering crisis. The entire physiology of the microcirculation in bone and bone marrow requires investigation, especially because the rheology of blood is affected by the presence of sickled cells and also by the numbers of white cells (I. Smaje, Wellcome Trust), which are often increased during crisis.

Although all patients with sickle cell anaemia have exactly the same genetic defect, there is wide variation in clinical expression of the disease. D. Weatherall (Oxford University) considered genes that modify the phenotype. Some of this

variation results from the coexistence of a deletion in another globin gene, α -thalassaemia (see Dover, G.J. *et al. Blood* **69**, 341; 1987). Thus a typical mendelian disorder has some of the features of a polygenic disorder. Other hypothetical compensatory mutations may affect the intraerythrocytic level of diphosphoglycerate (DPG), the main physiological regulator of the oxygen-haemoglobin dissociation curve. In other chronic anaemias, an increase in DPG levels improves oxygen delivery to tissues (Torrance, J. *et al. New Engl. J. Med.* **283**, 165; 1970); but in sickle cell anaemia this adaptive response backfires because the extra deoxygenation of haemoglobin favours sickling. Another component favouring sickling could be a pH-dependent ouabain-insensitive potassium channel that is apparently more active in the young red cells of sickle cell anaemia patients (C. Ellory, Oxford University). This channel may contribute to dehydration of red blood cells and the consequent promotion of irreversibly sickled cells. Both this potassium channel and the enzyme diphosphoglycerate mutase, which regulates the DPG level, are potential targets for drug therapy.

Treatment strategies can be broadly divided into five categories: anti-sickling agents; vasoactive drugs; promotion of persistence of fetal haemoglobin; bone marrow transplantation; and gene therapy. The first three approaches have failed so far, but there is increasing interest in the last two. Transplantation is available only to patients who have a suitable sib donor and the risks associated with the procedure and with graft-versus-host disease are still too high for this approach to become generally applicable in the treatment of sickle cell anaemia. But it is worth noting that bone marrow transplantation has been used successfully to treat about 200 children suffering from the more severe anaemia, thalassaemia major.

Gene therapy would avoid the problems associated with transplant rejection and graft-versus-host disease. Two approaches can be envisaged: replacing the mutated β -globin gene by homologous recombination; or insertion into the genome of haemopoietic cells of an additional copy of the normal gene. The ideal approach is the former, but it requires a substantial improvement in the isolation and culture of pluripotent stem cells as well as in the efficiency of homologous recombination. The second approach also requires technical advances in stem cell culture and the development of suitable

gene-transfer methods. In the meantime, the recent identification of a control element 60 kilobases upstream of the β -globin gene (F. Grosveld, National Institute for Medical Research, London) has solved the problem of obtaining high-level, tissue-specific expression of human β -globin in transgenic mice. Nevertheless, if gene therapy is going to be a cure for sickle cell anaemia (and many other diseases), a serious investment of resources must be made in this area.

Sickle cell anaemia has always been the paradigm for molecular analysis of genetic disease and the disappointing progress in treating it could be a warning for those studying other diseases. As is clear from the example of sickle cell anaemia, avoidance of this fate will require continued studies on the natural history of diseases and the physiological interactions in the whole organism: perhaps a true anti-sickling agent will then turn up.

The development of novel treatment strategies could be greatly facilitated by animal models, and it would seem to be a

high priority to adopt the new techniques of embryonic stem cell manipulation to create a mouse model of sickle cell anaemia (see Rubin, E.M. *et al. Am. J. hum. Genet.* **42**, 585; 1988). In some cases, mutations in the mouse models have failed to mimic directly the human disease. Notable examples include deficiency of the enzyme HPRT for Lesch-Nyhan disease and the *mdx* mouse models for Duchenne muscular dystrophy, respectively. But this serves only to emphasize the importance of physiological studies and provides an opportunity for discovering key steps in the disease process. Finally, if somatic gene therapy is going to be a viable approach for treatment of genetic disease, much more must be learned about the biology of stem cells, be it in muscle, brain or bone marrow. □

L. Luzzatto is in the Department of Haematology, Royal Postgraduate Medical School, Hammersmith Hospital, London W12 0NN; P. Goodfellow is at the Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK.

Geological boundaries

Extinctions and carbon cycling

M.L. Delaney

ARE mass-extinction events causes or consequences of changes in global carbon cycling? Reports on marine sequences by Holser *et al.*¹, Gruszczynski *et al.*² and Zachos *et al.*³ on pages 39, 64 and 61, respectively, of this issue conclude that substantial shifts in carbon partitioning among and within major sedimentary reservoirs occurred before, during and after extinction events defining the boundaries of the Mesozoic era. But the nature of carbon cycling associated with these mass extinctions, as recorded in the oceanic carbon isotope composition, is unique to each boundary. The Permian/Triassic mass extinction, about 250 million years ago (Myr), followed a prolonged interval of greatly increased net organic carbon oxidation. By contrast, the Cretaceous/Tertiary extinction (about 66 Myr) apparently resulted in the rapid cessation of oceanic primary productivity. After both boundaries, despite their different natures, recoveries of oceanic carbon isotope systems to more typical conditions took a substantial time.

Organic matter, formed in oceanic surface waters where light and nutrients are available for primary productivity, has a lower ratio (by about 2 per cent) of the rare, heavier stable carbon isotope (^{13}C) to the more abundant one (^{12}C) than the dissolved inorganic carbon. Therefore, dissolved inorganic carbon in surface waters, where organisms have used most of the available dissolved nutrients and

preferentially taken up the lighter carbon isotope, is relatively enriched in ^{13}C . The ^{13}C -depleted organic matter falls through the water column and is recycled back to its dissolved, inorganic constituents at depth, giving the dissolved inorganic carbon in deep waters a relatively low ratio of ^{13}C to ^{12}C . This surface-to-deep difference in the ratio of the two forms of carbon is thus generated by the cycle of primary production in surface waters and organic matter regeneration in deep waters. Consider a case where primary productivity in surface waters was greatly reduced (while nothing else about the ocean changed); $^{13}\text{C}/^{12}\text{C}$ ratios in surface and deep waters would become more similar, with the surface-to-deep gradient in carbon isotope ratio reduced.

Over longer timescales, the important carbon sources and sinks for the ocean as a whole are weathering of material from and burial of material into two sedimentary reservoirs: organic carbon, relatively depleted in the heavier isotope, and calcium carbonate, relatively enriched. Estimates are at present that the burial rate of carbon in the carbonate sink is four times that in the organic carbon sink. The average ratio of the two carbon isotopes in the ocean at any time reflects the balance of these geological cycles. For example, an increased supply to the ocean of carbon from organic carbon weathering, all else being equal, would result in a lower ratio of ^{13}C relative to the lighter

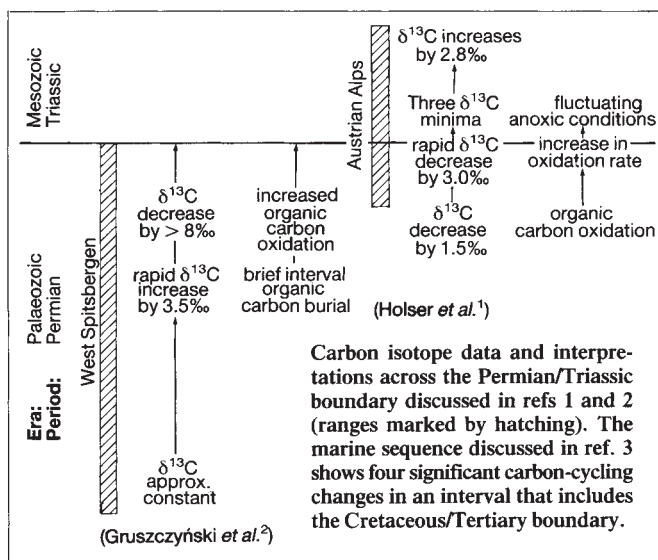
isotope in the ocean overall.

Surface-to-deep carbon isotope patterns can be studied by comparing the carbon isotope ratios in the calcite shells of surface and deep-dwelling organisms, as this material forms with little difference between its carbon isotope ratio and that of dissolved inorganic carbon. Long-term changes in carbon cycling affecting the ocean, such as decreases in organic carbon burial or increases in weathering, can be monitored by the ratio of the two carbon isotopes in shell material or in bulk carbonate sediments, reflecting the oceanic inventories of the two isotopes. Mass extinctions should result in dramatic decreases in primary productivity, and thus in carbon cycling within the ocean and relative to its sources and sinks⁴. The surface-to-deep gradient in ^{13}C to ^{12}C should be greatly reduced and the ratio of ^{13}C to ^{12}C in the ocean should decrease temporarily. The decreased surface-to-deep carbon isotope ratio has been previously noted for several Cretaceous/Tertiary boundary sequences, and a negative carbon shift has been observed for both of these extinction boundaries. The contributions¹⁻³ discussed in this article make clear the complexity of these transitions and the extended nature of the changes in oceanic carbon cycling, as recorded in carbon isotope records (see figure).

Holser *et al.*¹ present a detailed carbon isotope record from a continuous marine sequence in the Austrian Alps spanning the Permian/Triassic boundary, the most dramatic extinction event in the Phanerozoic. They conclude that this boundary was preceded by at least several million years of net oxidation of global reservoirs of organic carbon, and was followed by several anoxic episodes resulting from the long-term increase in the oxidation of organic carbon. The identification of these episodes of anoxia is supported by iron, sulphur and other minor-element data. But whether anoxia occurred globally or locally is unclear without data from other locations. The recovery to more typical carbon isotope values, and presumably to more normal oceanic oxidation-reduction conditions, took at least a few million years, indicating the complexity and long duration of this transition.

Gruszczynski *et al.*² use a carbon isotope record from brachiopod calcite from West Spitsbergen (present-day Arctic Ocean) to argue that the late Permian had a relatively brief interval of greatly increased organic-carbon burial before the start of the record presented by Holser *et al.*. The subsequent carbon isotope decline in their composite record is the largest yet observed in the late Permian. They suggest that the increased oxidation of organic carbon may have affected oceanic nutrient and atmospheric oxygen levels sufficiently to result in the mass extinctions.

Zachos *et al.*³ conclude that the



Cretaceous/Tertiary boundary was preceded by an interval of around 200,000 years of reduced net organic-carbon fluxes to the ocean. The boundary itself is marked by an extremely rapid cessation of oceanic primary productivity over less than 10,000 years, consistent with a bolide impact cause for these extinctions and apparently not related to the preceding shift in carbon isotopes. The drastic reduction in primary productivity persisted for at least half a million years at this low-latitude North Pacific site, with carbon isotope gradients between the surface and deep ocean not recovering to fully normal until about a million years after the boundary. This interpretation of these primary-productivity changes is enhanced by records of calcium carbonate accumulation and preservation and barium accumulation (as a proxy for organic carbon) at this site. Confirmation of the ocean-wide duration and intensity of these events requires more widespread geographic sampling of marine sequences spanning this boundary.

These three reports suggest intriguing links between carbon cycling and extinction events, with both events influenced by and resulting in changes in organic carbon cycling. Sea level seems to exert an important control on net rate of carbon oxidation and burial, with regressions favouring oxidation over burial^{1,3}. Better knowledge of the relationship between sea level and carbon cycling, along with unambiguous records of sea level, is required to document the impact of sea level changes.

There is a limit to what records of oceanic carbon isotope composition, used so effectively in these reports, can reveal about mechanisms and results of extinction events. As discussed above, the distribution of carbon between carbonate carbon and relatively ¹³C-depleted organic-carbon sedimentary reservoirs controls the average oceanic carbon isotope ratio, with the surface-to-deep gradient determined by organic-carbon formation and

regeneration. Dissolved phosphate, a critical, limiting nutrient supplied to the oceans by continental weathering, is incorporated in organic matter in primary productivity and recycled back to the dissolved state by organic matter regeneration in deep waters. Surface waters have relatively low dissolved phosphate concentrations, therefore, and deep waters are phosphate-enriched. Net organic-carbon burial in the sedimentary reservoir is an important sink for phosphate, estimated to account for 40 per cent of the total removal at present. Organic carbon is thus an important phosphorus carrier in the water column and an important removal mechanism for phosphorus from the oceans; any carbon cycling change has to affect oceanic phosphate inventories and distributions. For example, the marked decline in primary productivity inferred at the Cretaceous/Tertiary boundary should result in a more uniform water-column distribution of phosphate concentrations. The increased phosphate concentrations in surface waters, considered alone, should in turn result in increased rates of primary production, which would tend to re-establish

the water column phosphate gradient.

Also, the increased weathering of organic carbon before the Permian/Triassic boundary probably would have resulted in increased phosphate inputs to the ocean. These increased phosphate inputs should result in higher rates of primary productivity in the water column, if oceanic turnover times were similar. This in turn could result in higher rates of organic-carbon burial, negating the effects on the carbon isotope composition of increased weathering. Obviously, the mechanisms relating phosphate inventories, carbon isotopes and primary productivity are more complex than these simple examples. Carbon-isotope signatures alone cannot reveal linked changes in oceanic nutrients and primary productivity. Geochemical tracers indicating nutrient concentrations and productivity variations are needed to unravel oceanic carbon cycling changes around extinctions; Zachos *et al.*³ take a first step by using barium fluxes as indicative of organic-carbon fluxes. Ultimately, because organic-carbon weathering and burial influence atmospheric carbon dioxide and oxygen levels^{4,5}, an understanding of the links between these cycles will help to quantify atmospheric changes which can influence climatic changes. □

1. Holser, W.T. *et al.* *Nature* **337**, 39–44 (1989).
2. Gruszczynski, M. *et al.* *Nature* **337**, 64–68 (1989).
3. Zachos, J.C., Arthur, M.A. & Dean, W.E. *Nature* **337**, 61–64 (1989).
4. Tappan, H. *Palaeogeog. Palaeoclimatol. Palaeoecol.* **4**, 187–210 (1968).
5. Berner, R.A. *Am. J. Sci.* **287**, 177–196 (1987).

M.L. Delaney is in the Marine Sciences Committee, University of California, Santa Cruz, California 95064, USA.

RNA structure

Unwinding with a vengeance

Robert A. Lamb and Gideon Dreyfuss

UNTIL very recently, the old rule was that by examining the sequence of the genome template the RNA product could be predicted. The new rule is that the old rule is not quite good enough any more. In the past two years, a plethora of findings have underscored the tremendous plasticity of RNA molecules and the amazing variety of reactions they undergo. Recent examples include the two apolipoprotein B messenger RNAs, one of which has a uridine (U) residue instead of a templated cytidine (C)^{1,2}; the extensive addition of non-templated U residues to kinetoplast mitochondrial mRNA in trypanosomes³; and the addition of two non-templated guanosine (G) residues to a paramyxovirus mRNA⁴. Now, in the 23 December issue of *Cell*⁵, Bass and Weintraub report a striking addition to this new collection of modifications that change RNA structure. These investigators have discovered that a previously

known^{6–8} ATP-dependent unwinding activity also covalently modifies double-stranded RNA substrates by converting adenosine (A) to inosine (I) residues. This modification weakens the capacity for base-pairing and brings about essentially irreversible separation of the strands.

Other double-stranded (ds) RNA unwinding activities, both ATP-independent (for example, ref. 8; and several pre-mRNA proteins, see below) and ATP-dependent (such as the eukaryotic elongation factor eIF-4A, ref. 9) have been described in many organisms, but so far none of them has been shown also to involve modification of the RNA. As such, this novel activity described by Bass and Weintraub⁵ accomplishes the task with a vengeance, as the change to the structure and composition of RNA is both profound and irreversible.

The dsRNA unwinding/modifying activity was detected by careful but

serendipitous observations during studies on anti-sense RNA. What might be the normal purpose of this irreversible process? The most obvious answer might be an antiviral defence mechanism. All RNA viruses have to go through a dsRNA intermediate in their life cycle. Although it is well established that one general defence response triggered by the dsRNA form is the induction of interferon, the unwinding and irreversible modification of RNAs that are in a dsRNA form could complement the interferon effects. But does this in fact happen?

It is conceivable that the footprints of an antiviral dsRNA unwinding/modifying activity may have been left with unexpected and disastrous consequences, and may even have caused a human disease. Measles virus, a member of the paramyxovirus-morbillivirus family, has a negative-stranded RNA genome that is transcribed in the cytoplasm by the virion RNA transcriptase into a series of positive-strand mRNAs. In very rare cases of measles virus infection, a persistent infection occurs that is thought to involve defects in functions related to viral assembly, but not to RNA replication. This can lead to subacute sclerosing panencephalitis (SSPE), a progressive disease of the central nervous system.

Recently, Cattaneo *et al.*¹⁰ reported nucleotide-sequencing studies on cloned mRNAs isolated from autopsy material of SSPE brain. In one case they found that a discrete measles virus mRNA, whose protein product (M protein) is essential for virus assembly, has more than 50 per cent of its U residues changed to C. Although biased hypermutation caused by infidelity of the viral RNA transcriptase while transcribing only the M gene cannot be excluded as a mechanism for this¹⁰, it now seems plausible that the establishment of the persistent infection that led to SSPE results from the newly described dsRNA unwinding/modifying activity. If, during transcription of the viral genome, the M gene mRNA fails to leave the template, the dsRNA unwinding/modifying activity could change the A residues on the template RNA to I. In the next round of transcription of mRNA these I residues are copied to C residues. This two-step conversion mechanism could explain the extensive U to C changes found in the mRNA. It should be emphasized that this is not the sole explanation for persistent infections of measles virus. But the strong mutagenic potential of the dsRNA unwinding/modifying activity may have far-reaching consequences.

Another possible function of this novel activity could be its capacity to change specific RNA molecules, or groups of RNA molecules, including mRNAs. This could be accomplished by expression of complementary RNAs (such as small-nuclear RNAs) either to specific regions

within introns of pre-mRNAs, which could lead to alternative splicing patterns, or by modifications within the exons to bring about changes in the coding specificity. For example, a complementary small RNA that base-pairs to a region of an intron of a pre-mRNA which specifies its splicing pattern could stimulate the dsRNA unwinding/modifying activity to alter the relevant sequence such that it alters its normal splicing pattern. It is also possible that the activity found so far is merely one of many such activities and that other related ones with greater specificity lead to changes such as that found in the apolipoprotein B mRNA^{1,2}. Perhaps the activity is a general way of ensuring that segments of RNA that have the tendency to form double-stranded structures (intra- or inter-molecular) are maintained in single-stranded form. Should this not be reason for those in the field to look for inosine in their RNAs?

What could regulate dsRNA unwinding/modifying activity and prevent it from modifying most RNAs in the cell? Two general categories of regulation can be considered. First, there may be specific regulatory components that interact with it but which were separated from it during fractionation. Second, there probably exist activities that can counter the unwinding component of the activity, such as annealing-promoting proteins and proteins that block and protect regions of dsRNA from dissociation. A potent annealing activity has recently been described in nuclear extracts¹¹. Furthermore, RNA molecules exist in cells as RNA-protein complexes with RNP proteins. The recent finding of a complex assortment of RNA-binding proteins in pre-mRNA RNP complexes is an illustration of the abundance and diversity of RNP proteins¹². Many of them, being RNA-binding proteins¹², can function in unwinding and annealing of RNA molecules, and can bind to specific RNA sequences¹³, perhaps protecting them, and thus decisively influence RNA structure. Taken together with the sequence of the RNA itself, the interplay among all these factors will determine the structure and thereby the function of RNA in the cell. □

1. Chen, S.-H. *et al. Cell* **238**, 363-366 (1987).
2. Powell, L.M. *et al. Cell* **50**, 831-840 (1987).
3. Feagin, J.E. *et al. Cell* **53**, 413-422 (1988).
4. Thomas, S.M. *et al. Cell* **54**, 891-902 (1988).
5. Bass, B.L. & Weintraub, H. *Cell* **55**, 1089-1098 (1988).
6. Bass, B.L. & Weintraub, H. *Cell* **48**, 607-613 (1987).
7. Rebagliati, M.R. & Melton, D.A. *Cell* **48**, 599-605 (1987).
8. Wagner, R.W. & Nishikura, K. *Molec. cell. Biol.* **8**, 770-777 (1988).
9. Abramson, R.D. *et al. J. biol. Chem.* **262**, 3826-3832 (1987).
10. Cattaneo, R. *et al. Cell* **55**, 255-265 (1988).
11. Munroe, S.H. *EMBO J.* **7**, 2523-2532 (1988).
12. Piñol-Roma, S. *et al. Genes Dev.* **2**, 215-227 (1988).
13. Swanson, M.S. & Dreyfuss, G. *EMBO J.* **7**, 3519-3529 (1988).

Robert A. Lamb and Gideon Dreyfuss are in the Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208, USA.

Daedalus

Fictitious fat

As many would-be slimmers know, the body has a preferred weight which it defends stubbornly. Even if forced to unload fat, it does so in a specific sequence, which (as in the case of the dreaded 'pear-shape syndrome') may not correspond to the desires of the slimmer. An adult body has a fixed number of fat cells; a successful slimmer succeeds in emptying some of them, while others nearby remain full. This, says Daedalus, makes good chemical sense. A droplet of fat in a cell is potentially metastable. If metabolic demand begins to shrink it a bit, its radius will decrease, its internal pressure will rise, its free-energy balance will shift towards further shrinkage, and the process will accelerate until the droplet is completely metabolized. So a fat cell can have only two stable states, full and empty. It is in fact a digital information-storage element. And this, says Daedalus, is how the body knows how heavy it is. Even within a fat deposit, the cells fill up in a preferred sequence, and it need only record the position of the currently active cell to know how far the process has got. This 'full up to here' signal is presumably relayed by the nerve connection from the cell.

This theory explains the digital accuracy with which some people maintain a steady weight despite their chaotic eating habits. It also suggests a new, powerful, and effortless way to lose weight. You merely generate a false 'full up to here' signal, suggesting to the body that a specific fat deposit is far fuller than it really is. It will then shed fat from that deposit, in an effort to restore the proper apparent loading.

So DREADCO volunteers are being gorged on a wildly calorific diet, and subsequently starved back to normal weight, while an NMR tomographic-imaging system maps the preferred sequence of filling and emptying of their fat cells. Daedalus will thus locate the last cell to be filled in each fat deposit. A false 'full up to here' signal from this cell would imply that the deposit is utterly full. So mild electric shocks at this exact location should induce rapid and effortless slimming. Indeed, a commercial shock-slimming system, claimed to work by stimulating the muscles, probably owes its sporadic success to this effect. But DREADCO's version will launch its false signals with reliable accuracy, giving far more dramatic results. Conversely, a false signal launched from the first-filled cell of a fat deposit will cause that deposit to gain weight.

Thus DREADCO's Programmed Electric Body Sculpture will reshape the user's fat deposits under full control. The pear-shape syndrome will be banished forever!

David Jones

Seal vaccination success

SIR—In providing evidence that the primary cause of the epizootic of distemper with high mortality among harbour seals (*Phoca vitulina*) in the North and Baltic seas was canine distemper virus (CDV) or a closely related morbillivirus¹, which has since been claimed to be a new morbillivirus christened phocine distemper virus (PDV)^{2,3}, we heralded our evaluation in seals of a subunit CDV vaccine, shown to be effective in dogs⁴. If seals can be protected by the vaccine, it would not only confirm the close relatedness between PDV and CDV, but it would also provide final proof for this morbillivirus as the primary cause of the recent epizootic in seals. We now report our initial success with two different vaccines.

We used animals belonging to a breeding group of harbour seals kept in isolation for many years. So far no signs of phocid distemper have been observed in this group. For ethical and practical reasons the study was limited to eight adult animals. Six of these were vaccinated twice intramuscularly at two-weekly intervals followed by a third vaccination four weeks later. Two received a candidate subunit iscom CDV vaccine, prepared in collaboration with Coopers Animal Health Ltd (Berkhamsted, UK) as previously described⁴; the other four were vaccinated with a candidate inactivated whole virus CDV vaccine, kindly provided by Duphar BV (Weesp, The Netherlands). The other two animals were sham-vaccinated with an antigen-free preparation.

Before vaccination, CDV-neutralizing serum antibodies were undetectable (titre below 10) in these animals. After the third vaccination all six CDV-vaccinated animals had antibody titres ranging from 300 to 1,000. Ten days later all eight animals were transferred to a closed swimming pool and challenged, intraperitoneally or oculo-nasally, with a suspension of organ material from seals that had died during the outbreak.

Given similarly, this material produced clinical disease and the development of CDV-neutralizing serum antibodies in dogs⁵.

The two sham-vaccinated seals showed signs of acute respiratory disease and mucopurulent nasal discharge about two weeks after challenge and died on days 14 and 18. Viral antigen was detected in the spleens and/or lungs of both these animals with an enzyme-linked immunosorbent assay, using a monoclonal antibody which

recognizes the F protein of both CDV and PDV^{3,5}. No clinical signs were noted in any of the vaccinated animals (I. Visser, to be published). A threefold or higher (average 15-fold) increase in CDV-neutralizing serum antibody titres observed in all vaccinated animals two weeks after challenge suggests that some replication of PDV had taken place. As is often the case in the natural disease, CDV-neutralizing antibodies were not found in serum samples collected after the death of one of the two seals, probably because of the acute nature of the disease. The other had developed a CDV-neutralizing antibody titre of 30.

Although the number of animals studied is small, the data warrant the conclusion that seals can be protected from phocid distemper by vaccination with certain inactivated CDV vaccines. Since the disease was only produced in animals not

vaccinated against canine distemper, the last of Koch's postulates has also been fulfilled, reconfirming that infection with CDV or a closely related morbillivirus is the primary cause of the epizootic.

A.D.M.E. OSTERHAUS

F.G.C.M. UYTDEHAAG

National Institute of Public Health
and Environmental Protection,
Bilthoven, The Netherlands

I.K.G. VISSER

E.J. VEDDER

Seal Rehabilitation and Research
Center,

Pieterburen, The Netherlands

P.J.H. REIJNDERS

Research Institute for Nature
Management,
Texel, The Netherlands

J. KUIPER

H.N. BRUGGE

Ecomare,

Wadden and North Sea Center,
Texel, The Netherlands

Superconducting single crystals

SIR—Were large single crystals of the high-temperature superconductors of the type $\text{La}_{2-x}\text{A}_x\text{CuO}_4$ available (where A represents Ba or Sr), their physical and structural properties might provide insight into the still unknown mechanism of high-temperature superconductivity. By use of the travelling-solvent floating-zone method¹, we have grown such crystals and determined their superconducting properties, which we report. The anisotropy of their magnetic and electrical properties will be described elsewhere.

Appropriate amounts of La_2O_3 , CuO and SrCO_3 powders were mixed in ethanol, dried and calcined in air at 1123 K for 12 hours. The calcined powder was formed into a cylindrical shape, 6 mm in diameter by 50 mm in length, and compressed at a hydrostatic pressure of about 100 MPa. This rod was sintered at 1273–1473 K for 12 hours in oxygen gas and was then used as both the feed and the solvent rod. For crystal growth we used a double-ellipsoidal infrared heating furnace (Nichiden Machinery Ltd) heated by two 1.5-kW halogen lamps¹. Growth conditions

involved growth rates of 1.0–3.0 mm per hour, solvent compositions of 55–80 mol% CuO and a growth atmosphere of 0.2 MPa O_2 (to prevent vaporization of copper oxide from the melt zone).

A sample of $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ in this way, using a solvent of 75 mol% CuO, is shown in Fig. 1. The black, lustreless region at one end comprises a two-phase mixture of La_2O_3 and $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$, but the remaining portion, black with a metallic lustre, is single-phase $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. This region exhibits two facets parallel to the growth direction. Back-reflection Laue X-ray diffraction of these facets gives sharp diffraction spots, confirming the single-crystal nature, and reveals a fourfold rotation axis, identifying the crystallographic plane of the facets as (001). Although the samples contained some cracks, single crystals of about 6 mm in diameter by 15 mm in length could be obtained.

The lattice parameters of the tetragonal

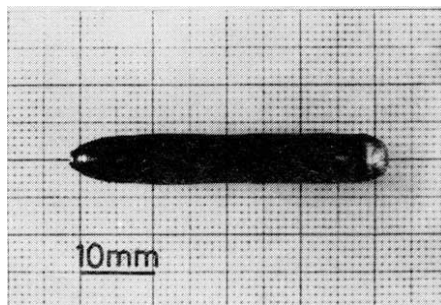


Fig. 1 Single crystal of $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. Dimensions are ~ 6 mm diameter and 40 mm length.

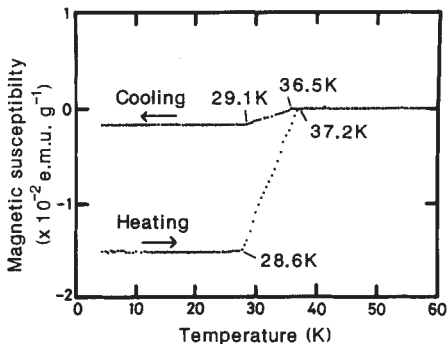


Fig. 2 Temperature-dependent magnetization of $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ single crystals. The upper curve is for cooling in a field of 0.936 Oe; the lower curve is for heating in 0.936 Oe after cooling in zero field.

- Osterhaus, A.D.M.E. & Vedder, E.J. *Nature* **335**, 20 (1988).
- Kennedy, S., Smyth, J.A., McCullough, S.J., Allan, G.M. & McNeill, F. *Nature* **335**, 404 (1988).
- Mahy, B.W.J. *et al.* *Nature* **336**, 115 (1988).
- De Vries, P., UytdeHaag, F.G.C.M. & Osterhaus, A.D.M.E. *J. gen. Virol.* **69**, 2071–2083 (1988).
- Osterhaus, A., *et al.* *Nature* **335**, 403 (1988).

$\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ crystals, measured by powder X-ray diffraction using Si as a standard, are $a=3.793\pm0.003$ Å and $c=13.20\pm0.02$ Å, which are consistent with values obtained from a sintered $\text{La}_{1.88}\text{Sr}_{0.12}\text{CuO}_4$ feed rod ($a=3.792\pm0.003$ Å, $c=13.21\pm0.02$ Å). Electron probe microanalysis indicates that the Sr content is uniform throughout the crystal, with $x=0.12$; for the feed material, $x=0.15$.

Figure 2 shows the magnetic susceptibility of $\text{La}_{1.88}\text{Sr}_{0.12}\text{CuO}_4$ as a function of temperature, measured using a SQUID magnetometer. The transition temperature observed on cooling the sample in a magnetic field of 0.936 Oe agrees well with that from heating the zero-field-cooled superconducting phase through the transition in the presence of the same

field; but the magnetization differs considerably between these two cases. The transition onset temperature and the end temperature are 37 K and 29 K respectively, so that the width of the transition, ΔT_c , is 8 K. This transition is considerably sharper than both that for sintered $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ (ref. 2), and that reported previously for a $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ single crystal. We found a Meissner effect of 13%.

I. TANAKA

H. KOJIMA

*Institute of Inorganic Synthesis,
Yamanashi University,
Miyamae 7, Kofu 400, Japan*

1. Kimura, S. & Shindo, I. *J. Cryst. Growth* **41**, 192–198 (1977).

2. Tarascon, J.M., Greene, L.H., McKinnon, W.R., Hull, G.W. & Geballe, T.H. *Science* **235**, 1373–1376 (1987).

Isotope dating of the Australian regolith

SIR—The work reported¹ by Bird and Chivas is based on three assumptions: (1) that materials recognized as forming deeply weathered profiles have been formed at some time in the past, on landscapes of low relief under tropical conditions; (2) that the date of such intense periods of weathering can be established by a method of sufficient accuracy to calibrate their oxygen isotope results; and (3) that the clay minerals used for oxygen-isotope analysis were formed during the intense weathering process.

We investigated the first assumption at two sites. The first site was the laterites and deeply weathered profiles on the Hawkesbury Sandstone near Sydney. These materials are generally regarded as erosional remnants of a fossil soil, formed in the Miocene and/or Pliocene, under intense tropical weathering, on a surface of low relief^{2,3}. Our mineralogical, geochemical and palaeomagnetic investigations^{4–6} demonstrate that these laterites and associated deep-weathered profiles were formed during the gradual stripping of the impervious cover of the Wianamatta shales, which provided the necessary environmental conditions for the mobilization of iron. In other words, all the laterites are the result of a time-transgressive process controlled by the rate at which the Wianamatta shale is being stripped from the Hawkesbury Sandstone, whereas the deep-weathering profiles are explicable in terms of normal bedrock variability.

Our second site was in the Cobar region of central western New South Wales, where silcretes and associated deep-weathered profiles had been recognized^{7,8} as having developed on a Tertiary pediplain. Detailed mineralogical and geochemical investigations of these materials^{6,9} show that the quartzitic Palaeozoic bedrock of these sites has remained essentially unaltered since a late Palaeozoic deformation and that both silcretes and deep-

weathered profiles can be explained in terms of normal bedrock variability.

Thus, even though the materials at these two sites have all the characteristics of deep-weathering profiles and associated laterites and/or silcretes, there is no necessary connection between such characteristics and their formation under some past tropical conditions.

Taking their second assumption, Bird and Chivas themselves¹ doubt the application of stratigraphic and palaeomagnetic methods. Yet they used these methods to calibrate the oxygen-isotope results. They do not comment on how the difficulties were dealt with to make the calibration meaningful. For example, in commenting on the use of stratigraphic methods, Bird and Chivas refer to Exon *et al.*¹⁰ to support the view that previous attempts to date Australian regolith have often been unsatisfactory. That paper, however, is only the last of a series which throws light not only on dating problems, but also on definitional problems.

The use of palaeomagnetism as a method of dating seems to be similarly flawed, for our investigation of iron in the Hawkesbury Sandstone^{5,6} suggests that no past intensive period of iron deposition coinciding with a period of deep weathering can be identified. What is indicated is that iron mobilization and deposition has occurred at a fairly uniform rate over the whole basin since the late Tertiary at least, and this is more than sufficient to explain any of the present-day iron accumulations. Our work casts doubt on studies that have established intense periods of iron deposition at some time in the Tertiary, and in particular on the determinations used by Bird and Chivas to establish the age of periods of deep weathering.

Finally, the long-term stability of the oxygen-isotope composition of kaolinites is fundamental to their use by Bird and Chivas in identifying the time of formation of clay minerals. If this is so, it can equally

be used to support the contention that in any weathering mass there is inevitably a contribution from inherited clay minerals, which, as we have shown for the materials derived from the Triassic Hawkesbury Sandstone and the Palaeozoic rocks of Cobar, can be the dominant factor. Even though the degree of inheritance is less in profiles developed on igneous rocks, inherited material will always be present. Bird and Chivas do not appear to have considered this factor.

These doubts about the application of the methodology to the dating of deeply weathered regolith lead us to question the significance of the three ages of weathering identified by Bird and Chivas. Such weaknesses must be resolved before any further progress can be made in the study of deeply weathered Australian regolith.

T.R. PATON

P.B. MITCHELL

P.A. HUNT

*School of Earth Sciences,
Macquarie University,
North Ryde, Sydney,
New South Wales 2109, Australia.*

BIRD AND CHIVAS REPLY—We fully agree with Paton *et al.* that the Australian regolith is a spatially and temporally complex entity, and that weathering remains a major surficial process on the Australian continent. Many of the points raised are valid, indeed obvious; but stringent space requirements precluded full discussion¹. We discuss these points in detail in a forthcoming paper¹¹, so here we will only correct some of the misconceptions held by Paton *et al.*

Their statement of the suppositions on which the study was based is erroneous. The supposition we make is that the isotope composition of meteoric waters in Australia has been changing with time.

We do not state that lateritization/deep-weathering requires tropical conditions and low-relief landscapes (although such conditions may aid the process); in fact, a major conclusion of this and our other studies¹² is that tropical conditions are not required. We do state that such weathering requires humid conditions and unimpeded drainage, and that as a result, infiltrating waters are not likely to have been subject to evaporative modification of their isotope composition.

We do not believe that such profiles can be independently dated with great accuracy (except in exceptional circumstances, such as dated weathered basalt overlain by basalt) which is why we have delineated only three age groups, spanning probably 200 million years.

We are aware that weathering can be an episodic or continual process, that weathering 'events' can be superimposed in a single profile, and that clay minerals can be inherited from the parent rock (although large quantities of kaolinite are

not common in most metamorphic or igneous rocks). But the increase in the isotope composition of meteoric waters with time means that low $\delta^{18}\text{O}$ clays (grossly out of equilibrium with modern waters) must have formed in ancient and cooler weathering regimes.

The division between fields B and C does not require the rejection of the independent age constraints, because in most cases the only independent constraint is simply that the weathering is older than mid-Tertiary. The division among pre-mid Tertiary samples is made on isotope grounds because the low $\delta^{18}\text{O}$ clays could not have formed during the comparatively warm climate of the late Mesozoic and early Tertiary in Australia. The isotope composition of the few samples of demonstrably pre-late Mesozoic age (such as kaolinite from regolith profiles beneath the Jurassic Wisanger Basalt, +10.2‰ and Jurassic Algebuckina Sandstone, +12.3‰) confirms the subdivision.

We are aware of the complexities involved in the palaeomagnetic dating of iron accumulations and, as indicated in our paper, we do not believe that there is necessarily a genetic or temporal link between the formation of a clay assemblage and an iron-rich zone in a regolith profile. We did not rely heavily on palaeomagnetically dated profiles for the reasons stated by Paton *et al.*, but in cases such as the Morney profile in south-west Queensland¹³, the palaeomagnetic date can be taken as a minimum age for the development of the clay mineral assemblage.

With reference to the work of Paton *et al.* on the Cobar region, note that a kaolinitic sample collected from a roadcut 23 km south of Enngonia (which did not have any independent age control) yielded a $\delta^{18}\text{O}$ value of +18.1‰ (post mid-Tertiary), consistent with the conclusions of those authors.

Finally, Paton *et al.* comment on the paucity of data points, particularly in the post-late Tertiary group, having previously described in detail the difficulties inherent in finding suitably unequivocal localities to sample. The post-late Tertiary samples we analysed were largely obtained from regolith developed on the late Tertiary–Quaternary volcanic provinces of Victoria and north Queensland. If Paton *et al.* wish to point out other such unequivocal localities, we would be glad to sample them.

MICHAEL I BIRD
ALLAN R CHIVAS

Research School of Earth Sciences,
Australian National University,
GPO Box 4, Canberra 2601, Australia

1. Bird, M.I. & Chivas, A.R. *Nature* **331**, 513–516 (1988).
2. Burgess, A. & Beadle, N.C.W. *Aust. J. Sci.* **15**, 170–171 (1953).
3. Faniran, A. *J. Geol. Soc. Aust.* **18**, 159–164 (1971).
4. Hunt, P.A., Mitchell, P.B. & Paton, T.R. *Geoderma* **19**, 105–121 (1977).

5. Bishop, P., Hunt, P.A. & Schmidt, P.W. *J. Geol. Soc. Aust.* **29**, 319–326 (1982).
6. Hunt, P.A. thesis, Macquarie Univ. (1985).
7. Langford-Smith, T. & Dury, G.H. *Am. J. Sci.* **263**, 170–190 (1965).
8. Dury, G.H. *J. Geol. Soc. Aust.* **13**, 299–307 (1966).
9. Wasson, R.J., Hunt, P.A. & Clarke, M.F. *Geoderma* **22**, 137–159 (1979).
10. Exon, N.F., Langford-Smith, T. & McDougall, I. *J. Geol. Soc. Aust.* **17**, 21–30 (1970).
11. Bird, M.I. & Chivas, A.R. *Geochim. cosmochim. Acta* (submitted).
12. Bird, M.I. & Chivas, A.R. *Chem. Geol.* **72**, 249–265 (1988).
13. Idnurm, M. & Senior, B.R. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **24**, 263–277 (1978).

Origin of life

SIR—Miller and Bada¹ have made a valuable contribution to the understanding of the organic chemistry of high-temperature hydrothermal systems. They are, however, incorrect in citing my work² as supporting the hypothesis that life began in vents at mid-ocean ridges³. This important hypothesis has stimulated much debate⁴ and is extremely interesting, but the evidence is strong^{2,5,6}, as Miller and Bada¹ have confirmed, that conditions are probably too extreme in submarine ridge vents.

The much more likely candidate for our birthplace is in subaerial (that is, land) hydrothermal systems. In them, temperatures at high level are commonly around 40 °C, with pH fluctuating, often even mildly alkaline. Closely juxtaposed parts of a system may vary greatly in temperature, and both water-dominated and relatively dry gas-dominated conditions are possible, often in rapid succession in the same place. There may be rapid influx of fluid of very different chemistry, from considerable depth. Many potential mineral catalysts occur. It is for these reasons that I concluded^{2,5,6} that the likely setting of an RNA world would be in a subaerial, meteoric water hydrothermal system. Perhaps nowhere else in the Archaean world were conditions so favourable for that extraordinary event, the origin of life.

E. G. NISBET

Department of Geological Sciences,
University of Saskatchewan,
Saskatoon S7N 0W0, Canada

1. Miller, S.L. & Bada, J.L. *Nature* **334**, 609–611 (1988).
2. Nisbet, E.G. *The Young Earth* (Allen and Unwin, Winchester, Massachusetts, 1987).
3. Corliss, J.B., Baross, J.A. & Hoffman, S.E. *Oceanologica Acta* **4**, Suppl. 59–69 (1981).
4. Joyce, G. *Nature* **334**, 564 (1988).
5. Nisbet, E.G. *Nature* **321**, 206 (1986).
6. Nisbet, E.G. *Episodes* **9**, 83–90 (1986).

MILLER AND BADA REPLY — Nisbet proposes that the origin of life took place in 40 °C vent waters rather than 350 °C vents. Russell *et al.*² have also claimed that lukewarm vent waters were the site of important prebiotic reactions. The vent waters that are referred to are apparently hydrothermal waters cooled by mixing with sea water or by loss of heat to rocks. Because the high-temperature conditions would decompose all the organic compounds,

the vent waters mixed with sea water would have less organic material than the ocean as a whole.

It is not clear what is to be gained from this process over heating parts of the primitive ocean to 40 °C on a beach or in a lagoon. An alternative is to heat the primitive sea water to 40 °C, but not much higher, by circulating it through heated rocks. This apparently does not occur near submarine vents, but may take place near volcanoes (H. Craig, personal communication). Again there seems little to be gained by invoking this type of process. The chemical reactions taking place at 40 °C are more rapid but still similar to those at 0 °C, with a few exceptions (for example, template polymerizations may not work at 40 °C because of the melting of the helix). Taking the heat of activation of a reaction as 20 kcal mol⁻¹ and assuming that 1 part in 10⁵ of the ocean is at 40 °C, then 99.9% of the reaction would occur at 0 °C and only 0.1% at 40 °C.

Russell *et al.* refer to the reported detection of glycine in Red Sea hydrothermal brines² as evidence that a biotic synthesis occurs in mild hydrothermal vents. This glycine is a contaminant, however, which bleeds from the Chexlex resin^{3,5} employed in the analytical scheme. In addition, no other amino acids were reported², which would be usual for a prebiotic synthesis.

If lukewarm vent water, and not the ocean as a whole, came in contact with minerals that catalysed critical prebiotic reactions, as proposed by both Nisbet and Russell *et al.*, our argument could change. It is not feasible to evaluate this possibility until the minerals and reactions are named. It is to be noted that an important role in prebiotic chemistry for minerals and clays has so far not been demonstrated. As for the role of hydrophobic surfaces a primitive oil slick⁶ may have been more important than the mineral surfaces suggested by Russell *et al.*

STANLEY L. MILLER

Department of Chemistry

JEFFREY L. BADA

Scripps Institution of Oceanography,
University of California, San Diego,
La Jolla, California 92093, USA

1. Russell, M. J., Hall, A. J., Cairns-Smith, A. G. & Brateman, P. S. *Nature* **336**, 117 (1988).
2. Ingmanson, D. E. & Dowler, M. J. *Nature* **286**, 51–52 (1980).
3. Lee, C. L. & Bada, J. L. *Earth planet. Sci. Lett.* **26**, 61–68 (1975).
4. Gardner, W. S. *Mar. Chem.* **6**, 15–26 (1978).
5. Bada, J. L., Hoopes, E. & Ho, M.-H. *Earth planet. Sci. Lett.* **58**, 276–284 (1982).
6. Lasaga, A. C., Holland, H. D. & Dwyer, M. J. *Science* **174**, 53–55 (1971).

Barnacle taxonomy

SIR—Crisp and Fogg¹ attack what they consider “mere changes of opinion” in establishing and changing names of taxa above the species level. To alleviate attendant inconveniences, at least among the barnacles, they offer a remedy of

compromise between the classification of Linnaeus, in which all species were lumped under *Lepas*, and a current system^{2,3}; namely, return to Darwin's classification of 1851–54 (ref. 4) and limit formal appellation of subsequently recognized species-group taxa to the subgeneric level.

It is possible, for some taxonomic reason or another, that a current genus might be more appropriately recognized as a subgenus. But it does not take a monkey wrench thrown into the gears of an entire system, as Crisp and Fogg have proposed, to enact such changes; when fine tuning is necessary, it can be made on an individual basis. No precedent would be set for, after all, "the despised" subgeneric category is still in use among the barnacles. Therefore I would like to offer the following rejoinder.

Darwin's attempts to estimate relationships between taxa, and concomitantly their patterns of diversity, were clearly hindered by taxonomic inadequacies and perceptions prevailing at the time. Many biological species were lumped as single species and, concomitantly, many now well recognized genera were lumped as single genera. Of *Balanus concavus* he wrote⁴, "... a good instance of the amount of variation which seems especially to occur in most of the species which have very extensive ranges." But this and the half dozen other 'variable' species he recognized have all turned out to represent several to numerous species^{5–7}. Thus, with their diversity and relationships poorly understood, it is no wonder Darwin found barnacle biogeography of "no particular interest, for the species are not sufficiently numerous; and, what is still more adverse, the genera, with unimportant exceptions, range all over the world; so that there is no scale of difference, and it cannot be said that these two regions differ in their species".

Thus knowledge of much of the diversity and our perception of the biogeography of barnacles languished for the more than a century after Darwin, despite the results of numerous expeditions and collections from deep as well as shallow water, and the advances in our understanding of the species concept and ecology. It took major systematic revisions to reach the point where barnacles became biogeographically interesting, and then to where they began not only to corroborate but to

elucidate general principles in marine biogeography^{8–11}.

What Crisp and Fogg fail to appreciate is that taxonomic recognition of the relationships between species and species groups, at and above the generic level, is necessary if evolutionary pathways and patterns are to be recognized. Continuing adjustments to such a system, usually enacted when knowledge has increased to the point where relationships are no

longer being accurately portrayed, are inevitable inconveniences. If someone eventually comes up with a better method of handling such a database, fine, but in the meantime the remedy of compromise proposed by Crisp and Fogg¹ does not satisfy these prerequisites.

W.A. NEWMAN

*Scripps Institution of Oceanography,
La Jolla,
California 92093, USA*

Antigen presentation by B cells

SIR—The recent letter of Lassila *et al.*¹ clearly demonstrates that B cells cannot be the antigen-presenting cells that lead to priming of helper T cells. This contrasts with data from several laboratories^{2–4} demonstrating that B cells are required for the priming of CD4⁺, class II MHC-restricted T cells that proliferate in response to antigen as discussed by De Franco⁵. We would like to offer the following explanation for these apparently conflicting results.

CD4 T cells belong to two separate subsets that differ in their functions^{6–9} and in their requirements for priming^{2–4,8–10} and accessory factors¹¹. One subset of CD4 T cells (helper T cells, or Th2) is effective in helping B cells make antibody, while the reciprocal subset of CD4 T cells (inflammatory T cells, Th1) is active in T-cell proliferation assays, cytotoxicity, macrophage activation, delayed type hypersensitivity and in fact will suppress B-cell responses^{12,13}.

Studies in B-cell-deficient mice indicate that B cells are required for the priming of the subset of CD4 T cells that proliferate in response to antigen^{2–4}, but are not required for the priming of the subset of CD4 T cells that help B cells secrete antibody^{8–10}. The data also demonstrate that the B cell required for priming the proliferating CD4 T cell must be antigen-specific and MHC matched to the CD4 T cell⁴.

By contrast, the clonal expansion of the Th2 subset requires IL-1^{10,11}, a molecule expressed abundantly on macrophage membranes but not on B cells or dendritic cells. Thus, the macrophage is almost certain to be the cell that presents antigen: self MHC complexes required for the priming of this CD4 T cell subset^{8,9}.

Given these facts, it is not surprising that in the experiment of Lassila *et al.* the subset of CD4 T cells that helps B cells could not be primed in the absence of syngeneic macrophages. However, this observation cannot be taken to mean that no priming of CD4 T cells has occurred. Had they measured priming of the CD4 T cell subset that participates in *in vitro* proliferative responses, one would expect such priming to have occurred. Likewise, the seemingly paradoxical observation that transgenic mice fail to make antibody

responses if they have B-cell but not macrophage class II MHC expression¹⁴ is probably due to the requirement for macrophages in helper T-cell priming. Finally, if Th1 tend to be suppressive of antibody production, the problem of avoiding the priming of helper T cells by B cell idiotypes, on which the Lassila *et al.* letter is based, may be moot.

The paper of Lassila *et al.* is an elegant demonstration of the fact that B cells are ineffective at priming those CD4 T cells that can help B cells. It will, however, require further study to determine whether this inability of B cells to prime CD4 helper T cells applies to all CD4 T cell subsets. Unlike the virgin CD4 T cells of Lassila *et al.*, we are certainly turned on by the idea that macrophages are required only to activate the subset of CD4 T cells that is most effective at helping B cells make antibody, while B cells almost certainly can prime those CD4 T cells that mediate inflammatory responses such as cytotoxicity, macrophage activation, or delayed-type hypersensitivity, as well as suppression of antibody responses^{9,12,13}.

KIM BOTTOMLY

CHARLES A. JANEWAY, JR

*Section of Immunobiology,
Yale University School of Medicine,
New Haven,
Connecticut 06510, USA*

1. Lassila, O., Vainio, O. & Matzinger, P. *Nature* **334**, 253 (1988).
2. Janeway, C.A., Jr., Ron, Y. & Sprent, J. *J. Immunol.* **138**, 2848 (1987).
3. Ron, Y. & Sprent, J. *J. Immunol.* **140**, 3773 (1988).
4. Kurt-Jones, E., *et al.* *J. Immunol.* **140**, 3773 (1988).
5. De Franco, A.L. *Nature* **334**, 199–200 (1988).
6. Kim, J. *et al.* *J. exp. Med.* **162**, 188 (1985).
7. Mosmann, T.R. *et al.* *J. Immunol.* **136**, 2348 (1986).
8. Janeway, C.A., Jr. *et al.* *Immunol. Rev.* **101**, 39, (1988).
9. Bottomly, K. *Immunol. Today* **9**, 268 (1988).
10. Bottomly, K. *et al.* *Eur. J. Immunol.* **10**, 159, (1980).
11. Greenbaum, L. *et al.* *J. Immunol.* **140**, 1555 (1988).
12. Bottomly, K. *et al.* *J. molec. cell. Immunol.* **1**, 42 (1983).
13. Tite, J.P. *et al.* *J. Immunol.* **139**, 2892 (1987).
14. van Ewijk, W. *et al.* *Cell* **53**, 357 (1988).

1. Crisp, D.J. & Fogg, G.E. *Nature* **335**, 120–121 (1988).
2. Newman, W.A. & Ross, A. *San Diego Soc. nat. Hist. Mem.* **9** (1976).
3. Southward, A.J. (ed.) *Crustacean Issues* **5** (1987).
4. Darwin, C. *A Monograph on the Sub-class Cirripedia* (Ray Soc., London, 1851 & 1854).
5. Ross, A. & Newman, W.A. *Trans. San Diego Soc. nat. Hist.* **17** (1973).
6. Dando, P.R. & Southward, A.J. *J. mar. Biol. Ass. UK* **60**, 787–837 (1980).
7. Newman, W.A. *Crustaceana* **43**, 25–36 (1982).
8. Southward, A.J. *Biol. J. Linn. Soc.* **20**, 59–72 (1983).
9. Newman, W.A. & Ross, A. *Antarctic Res. Ser.* **14** (1971).
10. Abele, L.G. *Biol. crustacea* **1**, 241–304 (1982).
11. Newman, W.A. & Foster, B.A. *Bull. mar. Sci.* **41**, 361–377 (1987).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references. □

Improving on Darwin

Niles Eldredge

Evolutionary Processes and Metaphors. Edited by Mae-Wan Ho and Sidney W. Fox. Wiley:1988. Pp.333. £35.95, \$85.

EVOLUTIONARY theory continues to be characterized by sharp dissent and often acrimonious debate. The 'synthesis' — that relatively tranquil period that began roughly in 1930 and ended in the mid-1960s — in retrospect seems anomalous. Since its birth in the 1850s, the subject has for the most part been a battleground; controversy is nothing new, but rather the normal state of affairs.

Many modern evolutionary biologists, myself among them, have proclaimed the dissension to be healthy. After all, the very essence of the scientific process entails questioning the received wisdom of the past. Yet at times the dialogue seems to transcend the simple search for a deeper, more accurate description of nature. Sometimes the theme is less one of correction and building upon a body of accumulated knowledge and more a clarion call to throw out the old — often with little clear development of what might take its place. The multi-authored contents of Ho and Fox's *Evolutionary Processes and Metaphors* tread a precarious path along the border between the bright search for a better understanding of evolution, and the darker quest to expunge modern biology of what are perceived as the evils built into Darwin's original ideas. The book's lowest moments come when darwinian thought is presented as at best an outmoded metaphor (shades of Jeremy Rifkin's *Algeny*) and at worst the personification of evil, as when competition rather than cooperation is seen to be central to a description of biological nature. This, and the repeated calls for a 'new paradigm', are the least prepossessing features of the book. How often one wishes that scientists would not take Thomas Kuhn so literally: no one wants to be guilty of doing mere 'normal science', so the search for new paradigms to overthrow the old becomes the norm.

But, for all its flaws, *Evolutionary Processes and Metaphors* is not just another exercise in Darwin-bashing — although those in whom the response is reflexive to defend reigning orthodoxy will have little difficulty finding enough absurdities and false characterizations of darwinian evolutionary biology to dismiss the book outright. If David Hull is right¹, and science (and indeed the entire world of ideas) proceeds in a fashion that basic-

ally mimics natural selection, more can be accomplished by dwelling on what seems to be promising in the book under review than complaining about the more flagrant red herrings it contains.

There are two basic sources of progress in scientific discourse: new data pertaining to phenomena previously unexplored; and conceptual clarification of phenomena previously considered, in some instances prompted by new data that reopen old issues. For example, several authors broach the topic of 'self-organization' in various biological systems, reflecting a resurgence of interest in this theme in evolutionary biology. And a number of the 15 essays touch on the recently reopened issue of Weismann's 'barrier'. Received wisdom, stemming from Weismann, held the distinction between germ line and soma to be inviolate, thereby ignoring the otherwise well-established facts that germ line is generated afresh from somatic tissue in plants, and is hardly 'sequestered' to the extent conventionally supposed in many animal phyla². The weismannian distinction of soma and germ line, with its asymmetric causal emphasis that the germ line determines the soma, but that what happens to the soma cannot affect the germ line, became the central dogma of molecular biology.

That somatic change can indeed become heritable naturally affects views on how the evolutionary process works. Thus I was particularly struck by Cullis's essay reviewing heritability of somatic mutations in flax, which poses the very real possibility that "higher plants have a genetically controlled variation generating system". Cullis takes pains not to claim that such induced variation arises as a response to organismic needs (a core neodarwinian postulate is that variation is generated without regard to organismic needs, that is, it is random with respect to selection). Yet in the very next essay, which is otherwise a balanced and useful account of the evolutionary implications of "new genetic mechanisms", Pollard asserts that the flax evidence indicates that germ-line DNA can develop "genetic variation in response to the modified condition" (emphasis mine) — thus, in nuance at least, going rather farther than Cullis in posing a more severe if apparently unsubstantiated challenge to orthodoxy.

Natural selection, perhaps unsurprisingly, provides much of the book's thematic continuity, as much in the guise

of 'metaphor' as 'evolutionary process'. Several authors question the propriety of the very term 'selection', as (they claim) nothing in nature is actually doing the selecting. It is, according to Ho (as but one example), as if natural selection had merely been substituted for Paley's Divine Watchmaker: Darwin fashioned natural selection as a mechanistic, naturalistic substitution for a concept of a super-natural Creator in order to explain organismic design in nature. While this is so, complaining that selection is an inappropriate term because there is no selector ignores the parallel with *artificial* selection that Darwin had in mind. It was simply a terminological parallel that led to the choice of word. The critique does not speak to the real issue: is there anything to the notion of natural selection as evolutionary process?

Some of the essays in the book, notably those by Ho, come close to concluding 'no'; because natural selection is an outmoded metaphor, it seems we are to conclude that it is not a process that occurs in nature, at least to any extent that would be helpful in understanding evolutionary history. In trying to emphasize the positive contributions of this book, though, let us consider a related theme — 'holism' — that seems to hold more promise. Some contributors complain that traditional selection theory treats organisms in a passive manner: organisms are commonly portrayed as being at the mercy of their environment. More strongly, a number of them feel that the organism has (in theory) been inappropriately divorced from the environment when actually organisms and environment form a holistic unity. Thus development, and phylogenetic changes in developmental patterns, can be seen as arising from interaction between organism and environment in a dynamic fashion that is not fully taken into account in conventional views of the relation between environment, adult phenotype and development — views in which natural selection plays, of course, an important role.

This line, when developed well (as in Oyama's essay) is thought provoking. Yet too often the conceptual lumping together of ontologically separate entities leads to confusion rather than clarity. The best example may be natural selection itself: Darwin saw that organisms are both economic and reproductive 'machines' (itself a loaded metaphor; in their introductory essay, Ho and Fox proclaim emancipation from just this sort of 'mechanical' world view). How well organisms within a local population fare economically, in a world of finite resources where not all offspring can survive to reproduce, is bound to affect their relative reproductive success. Yet the term 'fitness', defined holistically simply as probability of reproductive success, overlooks

1. Hull, D.L. *Science as a Process* (University of Chicago Press, 1988).
2. Buss, L.W. *The Evolution of Individuality* (Princeton University Press, 1988).

RESEARCH

NUTRITIONAL ECOLOGY OF THE RUMINANT

By PETER J. VAN SOEST. "A major synthesis in what has become an increasingly complex field."—*New Scientist*. \$35.20

INVERTEBRATE-MICROBIAL INTERACTIONS

Ingested Fungal Enzymes in Arthropod Biology

By MICHAEL M. MARTIN. Arthropods that eat wood, foliage, and detritus have difficulty digesting the cellulose in their food. This book describes the biological mechanism that allows some species to overcome this problem. 14 b&w photos, 11 drawings. \$16.45 paper, \$35.75 cloth

ZOOGEOGRAPHY OF CARIBBEAN INSECTS

Edited by JAMES K. LIEBHERR. This book unites for the first time the results of studies on a variety of insect groups native to the Caribbean islands, and is intended as an early progress report on the use of insects in biogeographical research from this area. 90 b&w photos, 86 drawings. \$43.95

THE NATURAL HISTORY AND BEHAVIOR OF NORTH AMERICAN BEEWOLVES

By HOWARD E. EVANS and KEVIN M. O'NEILL. This book brings together information, including much previously unpublished data, on the biology of more than two-thirds of the 34 North American species. 26 b&w photos, 30 drawings. \$24.75 paper, \$49.50 cloth

HAGAN AND BRUNER'S MICROBIOLOGY AND INFECTIOUS DISEASES OF DOMESTIC ANIMALS

Eighth Edition

By JOHN F. TIMONEY, JAMES H. GILLESPIE, FREDRIC W. SCOTT, and JEFFREY E. BARLOUGH. A new edition of a widely used classic. \$60.50

Cornell University Press

124 Roberts Place, Ithaca, NY 14850, USA
or: c/o Trevor Brown Associates
Suite 7, 26 Charing Cross Road
London WC2H 0LN

CORNELL

and obscures the various elements that may contribute to differential reproductive success. It is useful, in other words, to distinguish economic from reproductive activities within single organisms, the better to understand what causes affect reproductive success. Similarly, eliding the distinction between organism and (physical and biotic) environment threatens to obscure causal pathways, whether they be the role of environment in 'selection', or the effect organisms have on their 'environments'. No one denies that organisms and environments are intimately linked — no more than anyone denies that organisms are individuals, despite the usefulness of distinguishing their economic and reproductive functions when it comes to understanding the nature of natural selection.

Many essays in the book appear to be summaries of past work, in particular those devoted to the implications of phenomena newly apprehended. Especially at the molecular level, knowledge of such phenomena can only enrich our

understanding of evolutionary process. The assumption that molecular mechanisms that bias genetic transmission somehow call into question the validity of other biasing mechanisms posed in the past — natural selection, say — is fortunately only vaguely hinted at in some contributions. The essays I found most appealing (those by Hailman and Gray on methodological issues in teasing apart process from results in evolutionary studies; by Gordon and by Bateson on evolutionary aspects of behaviour; by Cullis on variation; and by Oyama on development) propose links to the darwinian paradigm on which they quite naturally wish to improve. Theirs seems the better grasp of the past; and they write with the implicit realization that, rhetoric aside, improvement in the conceptual basis of any science usually involves keeping what's right, fixing what's not and adding the apparent implications of newly gained knowledge. □

Niles Eldredge is Chairman of the Department of Invertebrates, American Museum of Natural History, New York, New York 10024, USA.

Highs and hopes

Roger Guillemin

Messengers of Paradise: Opiates and the Brain. By Charles F. Levinthal. *Anchor Press/Doubleday, New York:1988. Pp. 229. \$16.95.*

COMING after Abrams's *The Milk of Paradise* (1962) and Wright's *Dark Paradise* (1982), we now have Levinthal's *Messengers of Paradise*. Levinthal's purpose in this non-technical little book is to recount the story of the discovery in 1975 and 1976 of the enkephalins and endorphins, molecules found in the human brain (and in fact that of all vertebrates) which have the biological activities of opiates, and having done so to tell us how our own endorphins are responsible for the bits of paradise that occasionally come our way throughout life.

Levinthal, a professor of psychology at Hofstra University on Long Island, New York, first provides a short and entertaining history of the use of opium from antiquity to these days of massive illicit use. But why should the human brain have highly specialized receptors for the opiate molecules in the sap of the poppy? Because — so went the hypothesis — opium molecules possibly resemble some substances that are made normally by the brain and are involved in its dealing with those very same phenomena — pain and mood — that are modified by opiates. The next step was to prove it.

In yet another among many attempts to emulate Jim Watson's *The Double Helix*, Levinthal goes on to tell the story of the

search for these endogenous ligands of the opiate receptors, and their final characterization. But the difference between Watson's book, and all the others that have purported to describe the high adventure and competition involved in the making of scientific discoveries, is that Watson was centrally involved in the work that led to the proposal of a structure for DNA. He knew what he was talking about. His imitators have generally had only second- or even third-hand knowledge of what they have described, and the resulting books, although plausible to the lay public, have been deeply flawed.

Such is the problem with this part of *Messengers of Paradise*. I do not remember Levinthal's name in the literature dealing with the opioid peptides in the days when I was active in that field. Thus, the chapters that are meant to describe the intellectual process involved in the research, as well as the accompanying human drama, are unsatisfactory. Here readers may be easily entertained, but they will not become well-informed. This is not a work of scholarship. An accurate and complete historical study requires considerable time and devotion, and above all a scholarly sense of purpose in the endeavour (for a superb and rare example, see H. F. Judson's *The Eighth Day of Creation*, 1979).

So much for the messengers. What about paradise? Simple: paradise is a form of bliss, or vice versa, that is entered or glimpsed thanks to endorphins *à toutes les sauces*. Struggles over pain, social comfort, alienation and addiction, the quest for coherence and certainty, ecstasy, and finally creativity and laughter, are the subjects of a series of vignettes in which

everything is explained or proposed to be explained by too much, too little or just the right amounts of opioid peptides in our triune brains. After all, Somerset Maugham, in an exquisite essay on Zurbaran as a mystic painter (published in *The Vagrant Mood*, 1952), said plainly that mystical ecstasy, "generally the result of prayer and mortification", "may also arise through the influence of a drug, opium, for instance".

I have a lot of sympathy for some of the ideas and proposals described by Levinthal, and nothing would please me more than to be convinced that they are substantiated or even substantiable. There is little doubt that the opioid peptides are involved, somehow, in our homeostatic handling of a variety of painful events. But even here I say *somehow*, as there are still many unanswered questions. For instance, I do not know of a definitive demonstration that acupuncture 'works' exclusively through the endorphin system. The 'proof' in such statements is usually that the expected effects are allegedly prevented by prior administration of naloxone, a powerful opiate antagonist. In the hands of critical investigators, such studies are difficult to perform and even more difficult to control. Moreover, we now know of effects of opioid peptides which are not naloxone-reversible.

A characteristic of many of these facile interpretations is the persistent confusion between the significance of blood levels of endorphins (which represent levels of endorphins of pituitary origin) and what happens or may happen in brain levels of endorphins, in one system of neurons or another. We have, so far, practically no way of knowing what is going on in such systems, even in laboratory experiments let alone the clinic. Intravenous injection of massive amounts of β -endorphin into animals or man has no effect whatsoever on measured functions of the central nervous system, because peptide molecules of this type do not penetrate the blood-brain barrier. Thus, relating the long distance runners' 'highs' to their blood levels of endorphins is meaningless.

Yes, the discovery of the chemical structure of the endogenous opioids found in our brains was a remarkable achievement. But it is still too early to know what to make of it, as basic knowledge or in daily life. Most of the original questions about the role(s) in health or disease of the enkephalins, the endorphins, and the other opioid peptides that were recognized later, still remain wide open. For all I know, some of Levinthal's enthusiasm — and, let's face it, that of most of us originally involved in these discoveries — may still be vindicated. □

Roger Guillemin is Research Professor and Chairman of the Laboratories for Neuroendocrinology, Salk Institute, La Jolla, California 92037, USA.

A growing sphere of knowledge

Ghilleen T. Prance

Biodiversity. Edited by E. O. Wilson. National Academy Press: 1988. Pp.521. Pbk \$19.50. Distributed in Europe by Wiley, £16.80. To be published in hard-back next month, £27.85, \$32.50.

UNTIL recently, 'biodiversity' — short for biological diversity — was a term used only by a few biologists. Today it is the subject of bills in the US Congress and of US Aid funding, and it was the central concern of a special meeting of the Pontifical Academy. Biodiversity has become the dominating theme of many scientific conferences, as it was at the August 1988 meeting of the American Institute of Biological Sciences held in Davis, California.

There is good reason for this rapid upturn of interest in the topic. In the past two years, alarming details have emerged of the rates of destruction of the tropical rainforest that harbours some 50 per cent of the world's biological diversity. Satellite photographs have shown that, in 1987, 204,000 km² of forest were burned in Brazil, and for 1988 the estimate is near to 250,000 km², much of the loss being of primary rainforest. A few years ago, environmentalists were called to task for suggesting that the annual worldwide total of rainforest destruction was 50 km². But the proven rate of destruction, the patent effect on climate and the number of species that are in danger of extinction have now combined to make biodiversity the centre of scientific and political attention.

Biodiversity, the book under review, is based on a national forum which took place in September 1986 at the Smithsonian Institution, under the auspices of the US National Research Council. Made

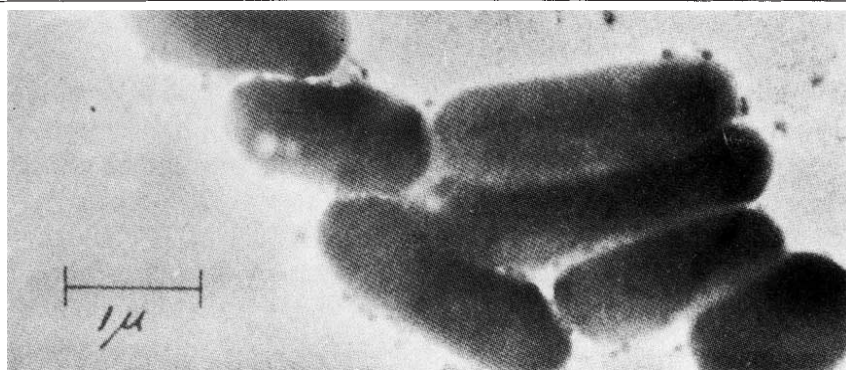
up of 57 chapters written by specialists, it is a thorough review of the many aspects of the subject. Although the emphasis is on tropical rainforest, the threats to many other ecosystems, such as grasslands, oceanic islands, coastal zones and oceans are covered. Two of the main facets of biodiversity are stressed — species diversity, and habitat and ecological diversity. Genetic diversity is not considered in any detail.

There is a particularly useful section of five chapters on the value of biodiversity, and another on the contribution science and technology can make. Here, the role of botanical gardens and zoos is given prominence. Other topics covered include agroforestry and other sustainable systems of management, political aspects of protecting diversity and ethical issues. The contents of the final section, "Ways of Seeing the Biosphere", vary from a native American account from New Mexico giving the viewpoint of a Pueblo Indian, to a description of the Earth as a living organism by James Lovelock, the creator of the Gaia hypothesis.

Each contribution is brief and to the point, the most of them have references to further work on the particular subject dealt with. The book does not contain many new ideas or concepts; rather, its value lies in bringing together an appropriate diversity of human talent to address a topic that is vital for the survival of life on Earth. It is written in such a way as to be understood by the general reader and, with its broad scientific content, will be used for teaching courses.

In the two years since the national forum was held, immense amounts of further evidence of the threat to biodiversity have been gathered. Yet with its comprehensive treatment the appearance of this book remains timely. It deserves wide notice. □

Ghilleen T. Prance is Director of the Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB, UK, and a visiting Professor at the University of Reading.



One of the earliest electron micrographs, taken by T.F. Anderson and Max Delbrück in 1941–1942, of bacteriophage on the surface of *E. coli*. It showed (amongst other things) that phage never enter the cell, an observation made earlier and independently by H. Ruska in Germany. The picture is reproduced from *Thinking about Science: Max Delbrück and the Origins of Molecular Biology*, by Ernst Peter Fischer and Carol Lipson (W.W. Norton, \$19.95; to be published in Britain on 18 January, £13.95). The book is a translation and revision of Fischer's *Licht und Leben*, reviewed by Max Perutz in *Nature* 320, 639 (1986).

A walk on the wilder side

A. Rupert Hall

Let Newton Be! A New Perspective on His Life and Works. Edited by John Fauvel, Raymond Flood, Michael Shortland and Robin Wilson. Oxford University Press: 1988. Pp.269. £17.50.

Isaac Newton, an eccentric professor of mathematics at Cambridge unaccountably chosen for the chief office of the Royal Mint in 1696, was author of two books, published posthumously, *The Chronology of Ancient Kingdoms Amended* and the *Prophecies of Daniel and John*. Further, he left behind a vast quantity of manuscripts dealing with his business at the Mint, his views on the doctrine of the early Church, and alchemy. As scholar, theologian and monetary expert Newton was neither original nor up to date, and in these fields his work has been passed by without ever receiving much consideration.

If this had been Newton's obituary, it is doubtful that he would be of more interest today than Archbishop Ussher, even to scholars. But in fact Newton also wrote the *Principia*, *Opticks* and *Optical Lectures*, a large number of papers in *Philosophical Transactions* and several mathematical texts. It is a truism to affirm that if he has any place in the history of civilization, it is because of his scientific writings.

This elegantly produced collection of papers — with an appalling dust jacket — tells us more about the eccentric don than about the experimenter and mathematician known to scientists. It is all, or nearly all, quite true: Newton (like other mathematicians) did look into 'equal temperament' in music and he did have (to us) bizarre ideas about the reappearance of the harmonic ratios in other parts of physics, such as the optical spectrum (as Penelope Gouk relates). He did (as P.M. Rattansi and she report) draft learned pages on the Pythagoreans' ancient knowledge of the deeper principles of physics, such as the inverse-square law of gravitation. He did meticulously ascertain the life-span of the locust (species unstated) in order to calculate the prophesied length of the expansion of Islam. And he did (as J. Henry and J. Golinski explain) painstakingly copy out the (apparent?) gibberish of the alchemists, with their fanciful symbolic pictures, and himself experiment on the volatility of metallic salts and the reality of Diana's Doves.

As Keynes wrote long ago, Newton's mind was complex and strange. He accepted as unquestionable the priority of Hebrew learning. He was equally confident that ancient texts were pregnant with metaphors and codes, whose meaning could not appear to the unlearned.

Only by discovering it, and so disclosing and restoring the pure wisdom of the infant world, could men perform their duty to God by understanding his Creation and Providence and so prepare for his Son's second advent. As P.M. Rattansi says, natural knowledge was but one element in Newton's *weltbild*, of whose theological character a strong hint is conveyed in the final scholium of the second *Principia* (1713).

Since Keynes (and so many others afterwards — nothing here is really new) read the pages that Newton would have kept hid, our knowledge of him has become richer but more confusing. Scholars are in danger of writing about the Newton who was insignificant to the neglect of the Newton who transformed thought. To be dismissive of Newton — as Whiteside, Westfall and Cohen, for example, have *not* been — is absurd. We see in this book two pictures of Harpo Marx representing Isaac Newton in the 1957 film *The Story of Mankind*, and many images taken from the books of the mystic

'platonist', Robert Fludd; it is doubtful if any of these is relevant to either Newton, the open or the concealed. Nor was Newton himself muddled about the natures of the diverse studies he pursued with equal ardour.

We need to be cautious about the inference that the scientific Newton drew inspiration and principles from the esoteric reading of his other self. My impression is that when experimenting or measuring or developing a structure of geometrical theorems, Newton was as much a 'scientist' as if he had had no other interests. In this book the investigations of the scientific Newton are well described, though necessarily in elementary language, by John Roche (*Principia*), C. Hakfoort (optics) and Jon Pepper (mathematics) — no place here for the kind of technical article ("Newton and the Calculus of Variations") that seemed appropriate 60 years ago. □

A. Rupert Hall, 14 Ball Lane, Tackley, Oxford OX5 3AG, UK, is Emeritus Professor of the history of science and technology, University of London.

Pollen stations

John Birks

Paleopalynology. By Alfred Traverse. Unwin Hyman: 1988. Pp.600. Hbk £50, \$75; pbk £24.95, \$34.95.

PALAEOPALYNOLOGY is the study of fossil pollen grains, spores and other organic microfossils (excluding diatoms and nannofossils) preserved in rocks and sediments. Among the objects of interest are acritarchs, dinoflagellates, microforaminiferal inner tests, chitinozoans, scolecodonts, selected algae, fungal, bryophyte and pteridophyte spores, and, of course, pollen from the recent past to the Precambrian. With such an enormous taxonomic diversity and stratigraphical range, it is inevitable that palaeopalynologists tend to be divided according to their specific stratigraphical or taxonomic interests, or both. And it is not surprising that, until now, nobody has written a book in English that sweeps across the subject as a whole.

Alfred Traverse, of Pennsylvania State University, has done just that. His book covers what palaeopalynology is and why it works; the basic biology and chemistry of palynomorphs; the stratigraphical palynology of the Precambrian and all the way through to the Quaternary; the production, dispersal and deposition of modern pollen; and applied palaeopalynology and laboratory techniques. It is profusely illustrated and has an extensive and up-to-date bibliography. Despite the enormous amount of information included, it is easy to read, with fascinating comments about many of the pioneers of the subject.

The book's greatest strength lies in the wealth of practical information it contains — the appendix on laboratory procedures is a monograph in itself and provides palynologists with an invaluable illustrated account of preparation techniques. The chapter on pollen morphology abounds in common sense and sound judgement, and includes a helpful synthesis of the different morphological terminologies and philosophies that are involved in understanding pollen-wall structure, along with an excellent account of the neglected topic of pollen orientation. The 30-page glossary is important for everyone concerned about precise terminology.

Traverse has maintained an admirable balance between the chapters on stratigraphical palynology and has avoided the temptation to put undue emphasis on the Tertiary (his main speciality). The book's weaknesses are the inevitable result of its breadth — some topics seem to my Quaternary palynological eyes to be covered too superficially. For example, there is virtually no mention of pollen-analytical studies of the effect of prehistoric man on vegetation, one of palynology's great success stories in Europe; and little is made of palynology's contribution to climatic reconstructions in the Quaternary or to modern plant ecology.

Traverse's book contains much of interest to all palaeopalynologists and it will, I hope, help to narrow the gaps between the different sub-disciplines. I urge all pollen analysts, novice or expert, to acquire and read it, and to share the author's manifest enthusiasm for his subject. □

John Birks is a Professor in the Botanical Institute, University of Bergen, Allegaten 41, N-5007 Bergen, Norway.

Eighteen eighty nine and all that

W. F. Bynum and J. L. Heilbron

An anniversarial cornucopia featuring rubber gloves, Louis Ilosvay de Nagy-Ilosva, the Eiffel Tower, nuclear fission, the Swedish Academy of Sciences, J. Willard Gibbs, the omega minus meson, Antoine Laurent Lavoisier, Timothy the tortoise and much, much more.

IN continuance of the public service begun here last January, we again fill the slate of impending anniversaries according to the decimal system. In this we have subordinated our own preference, which runs to base 60, to the wishes of our many correspondents. The century remains our unit; boldface numbers in parentheses indicate multiples of a hundred years. The progress of science cannot be arrested for long, however. We observe, as a hint for the future, that 'hundred' has signified 106 in counting sheep, 120 in counting balks, boards and bomspars, 124 in counting cod or haberdine, and 225 in counting onions and garlic.

In the mundane reckoning, the French Revolution (**2.0**) may occupy pride of place. The storming of the Bastille on 14 July 1789 was greeted variously throughout Europe and the New World, but time has softened the passions and mellowed the emotions. We can hope that *liberté, égalité, fraternité* will be practised as well as celebrated in 1989. Science and technology, too, can claim their revolutions, both intellectual and social. Among the biggest of conceptual revolutions was that in chemistry, which flowered in the city of the Bastille in the year of its storming; Lavoisier, the prime mover in the one revolution, became a victim of the other. Among the social-technological revolutions on which we can reflect in 1989 are those founded on photography, nuclear fission, the jet engine and miracle drugs.

1889 (1.0 centenary)

This year, in Spring, admirers of gigantism have to celebrate the hundredth anniversary of the completion of the Eiffel Tower. Built for the international exposition of 1889, the tower at 985 feet rose twice as tall as the tallest structure then in existence; it stood and stands for the height of wrought-iron technology. It has also been an instrument of scientific experimentation. In 1909 (**0.8**) Thomas Wulf designed a quartz-fibre electroscope especially suited to detect the presence of

ionizing radiations. The following year he carried it to the top of the Eiffel Tower and found that its rate of discharge did not decline with height as expected. Confirmation and extension of this result during balloon flights by Victor Hess and others led to the discovery of cosmic rays. The

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Towering achievement — a contemporary Punch caricature of Alexandre Gustave Eiffel (Mary Evans Picture Library).

underlying assumption of the measurement, that the leak of an insulated body exposed to the air does not in general arise from failure of the insulation, had been demonstrated definitively 20 years before (**1.0**) by C. V. Boys.

Shortly after the Eiffel Tower peaked out at 300 metres, Ludwig Mond — born 1839 (**1.5**), died 1909 (**0.8**) — found a cheap way to make replicas of it in nickel. This was the discovery of nickel carbonyl, which gave wings to the heavy metals (as Kelvin put it) and also much profit to Mond enterprises, because it led to a cheap way to extract nickel from poor ores. The profit in turn gave wings to

British scientific and cultural life through Mond's benefactions, which included gifts to the National Gallery, the Royal Institution and the Royal Society.

Further towards higher things, 1889 was a fine year for students of the stars. Edward Charles Pickering (died 1919 — **0.7**), a passionate photographer of the heavens above Harvard, discovered the first known spectroscopic binary — a double star not resolvable by the telescope, but detectable by the periodic doubling of its spectral lines. Astronomers then residing on Mercury and facing the Sun would not have discovered this interesting phenomenon, nor, indeed, would they ever see any star but ours; for, as Giovanni Virginio Schiaparelli observed 100 years ago, Mercury constantly keeps the same surface towards the Sun. One of us insists on adding here that the total solar eclipse of New Year's day, 1889, was visible in California, and that a month later gratified observers founded the Astronomical Society of the Pacific. The total solar eclipse visible in Guiana just before Christmas does not seem to have inspired a similar organization.

With all this we have not mentioned the items of 1889 that showed science at its most promising, to one observer at least. In 1917, a Hungarian writer, Louis Ilosvay de Nagy-Ilosva, predicted a glorious development of science after the Great War. He

rested his case for creative post-war scientific activity on a list of discoveries made after the Franco-Prussian War. For 1889 he gave the invention of organotherapy by Charles-Édouard Brown-Séquard (testicular injections which B-S believed had rejuvenated him); the demonstration by Nathan Zuntz that albumin, carbohydrates and fats are equally usable by the muscles; and the foundation by Albert I of the oceanographic museum of Monaco.

It might seem trivial to celebrate so humble a thing as the rubber glove, but modern surgery would not be the same without it. Rubber gloves had been possible since the vulcanization of rubber

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

"Jabez Hogg and Mr Johnson", a daguerreotype taken in c. 1843. The picture is the earliest known of a photographer at work. Hogg is timing the exposure with his watch, holding the lens cap in his other hand; the head of Mr Johnson is supported by a rest, to help him to keep still for an exposure lasting as long as a minute. It is not known what part the caged bird had in the proceedings.

became practicable from 1839 (1.5); indeed various surgeons and accoucheurs had advocated their use thereafter, generally to protect themselves from nasty diseases from which their patients might be suffering. But the continuous history of the surgical glove dates from 1889, when W. S. Halsted asked the Goodyear Company to make two pairs, with gauntlets, for his scrub nurse, Miss Caroline Hampton (who married him the following year). Their original purpose was to protect Miss Hampton's hands from the disinfectant to which she was allergic, but the wearing of gloves gradually caught on for surgeons as well.

Halsted was the first professor of surgery at the Johns Hopkins Hospital, which opened in 1889 in anticipation of the Johns Hopkins School of Medicine (1893). Theirs will be a lengthy celebration.

No centennial list would be complete without mention of the transmigration of souls. In 1889 there died Michel Eugène Chevreul, aged 103, Warren de la Rue, James Prescott Joule and Maria Mitchell. The same year saw the births of Sewell Wright (who lived almost to witness his own centennial), Lord Adrian, Dirk Coster, Ralph Fowler, Beno Gutenberg, Edwin Hubble, Ludwig Wittgenstein and the historian of science Hélène Metzger. We have not been able to discover with certainty which souls went where.

1839 (1.5 centenary)

Let us hope that the productive mixture of science and art that gave us photography recurs in celebrations of the sesquicentennial of its invention. In August of 1839

Louis Jacques Mandé Daguerre — born 1789 (2.0) — painter of panoramas, proprietor of the Paris and London Dioramas, and careful chemist, announced his process for sensitizing metal plates, exposing them to light and fixing the latent image. The same year, William Henry Fox Talbot, Cambridge wrangler, philologist and frustrated landscape painter, made known his process for taking negative images on treated paper. Many far-reaching improvements perfected in the next half-century made photography an indispensable tool of science. By 1889 (1.0) the physiologist Étienne-Jules Marey could determine with its aid that a fly in flight flaps its wings 288 times a second, a busy bee only 190 times, a lethargic butterfly 9 times. And the physicist Friedrich Neeson could propose to investigate the details of a more sinister sort of flight, the speed and spin of artillery shells.

American physicists will want to do honour to their first great theorist, J. Willard Gibbs, born 11 February 1839 in New Haven, Connecticut, where he spent all his life after returning in 1869 (1.2) from three years of study in Europe. He made himself known by sending his precise and difficult memoirs on thermodynamics and statistical mechanics to the leading European physical scientists. Maxwell, Rayleigh and Ostwald read and admired: Helmholtz, Planck and Einstein read not, and had to recover Gibbs's results independently, and with difficulty. Gibbs's years in Europe overlapped those of his exact contemporary, Aleksandr Grigorievich Stoletov, who became professor of physics at the University of

Moscow. His best-known work, which established certain regularities of the then-new photoelectric effect, is eligible for centennial observance.

Prophetic observers in 1839 might have read into the official opening of the Physiological Institute at the University of Breslau something of moment. At the least, they would have known that it was the first physiological institute within the burgeoning German university system and that its director and leading light, Jan Evangelista Purkyně (died 1869, 1.2), had already made a number of important physiological and histological discoveries. Like any serious microscopist of his day, Purkyně had availed himself of the improved accuracy of the achromatic microscope: "With boundless eagerness I investigated within the shortest time all areas of plant and animal histology and concluded that this new field was inexhaustible".

After 1839, Purkyně could make use of a systematic theory of cells, enunciated in that year by Theodor Schwann, whose *Mikroskopische Untersuchungen* capped what had been five amazing years of experimental productivity. Schwann's cell theory is not our cell theory, for he postulated that these basic units of living organisms can be formed by a process analogous to crystallization from a non-living soup he called the *blastema*. *Omnis cellula e cellula* came a little bit later, and not from Schwann. He himself returned to the Roman Catholicism of his fathers, discovered the "God of the heart, not of reason", and sought enlightenment through mystical insights into the nature of things. Appropriately the Microscopical Society of London (now graced with Royal blessing) was founded in that year of microscopical revelation.

Geologists will want to note the death, on 28 August 1839, of William Smith, the surveyor and stratigrapher whose maps did so much to establish historical geology and palaeontology. Like virtually all geologists in the discipline's heroic age, Smith travelled indefatigably, although confining himself to the British Isles. Sir Roderick Impey Murchison's *Silurian System* of the same year limited itself to Wales and the west of England, although Murchison possessively investigated Silurian remains in Russia and elsewhere. Another travelling naturalist, Charles Darwin, also burst upon the Victorian reading public with his *Journal of Researches*, though with little hint of the ideas that were later to revolutionize the study of the natural world.

Our preoccupation with the health of the ozone layer may promote a sesquicentennial for Christian Friedrich Schönbein — born 1799 (1.9) — who in 1839 identified the origin of the odour around electrostatic machines and electrolytic oxygen as a new sort of gas. He

was then professor of physics and chemistry at the University of Basel, where he arrived following a curriculum vitae that included a strong dose of *Naturphilosophie* and an even stronger antidote, in the form of a year's stint as mathematics master at a boys' school in Epsom, Surrey. It took over a decade for Schönbein's colleagues to convince him that ozone is allotropic oxygen, and not, as he insisted, a chemical compound. As our next entry further illustrates, oxygen has not been an easy concept for its discoverers.

1789 (2.0 centenary)

Enlightened people wanted not only to talk well but also to talk correctly. French chemists shared and profited from this compulsion. In 1789 Antoine Laurent Lavoisier gave them a *Traité élémentaire de chimie* incorporating a nomenclature that was, in the buzz words of the Enlightenment, reasonable, natural and systematic. The revolution worked by his systematic vocabulary and the rules for applying it may best be illustrated by the substitution of words understandable by a modern chemist for terms now unintelligible, such as flowers of zinc and butter of arsenic. But just as even the most punctilious lexicographer sometimes fumbles etymologies, so Lavoisier invented some words that have proved unsuitable. Of these, 'oxygen', coined in the belief that dephlogisticated air is the principle of all acids, is perhaps the most interesting. Although Humphry Davy proved the etymology false in 1809–1810 (1.8) by demonstrating that muriatic acid (HCl) contains no oxygen, the misnomer survives. No one now has any difficulty with it; it is not necessary to know what a word means to use it correctly.

Another triumph of the natural and the systematic in 1789 was the publication of Antoine-Laurent de Jussieu's *Genera plantarum*, which presented botanists with an improved taxonomic system developed by the inbreeding of the ideas of three generations of Jussieus. This classification of flowering plants rested on correlation of a wide variety of characters and consideration of a very large number of specimens. Like Lavoisier's *Traité élémentaire*, Jussieu's *Genera plantarum* quickly became the cicerone in its field; although it has suffered many changes in its lower taxa, its arrangement of families has remained a part of botanical classification.

Jussieu was part of the increasingly professional tradition of French science, the Reverend Gilbert White a quintessential part of the amateur one in Britain. His book, *The Natural History and Antiquities of Selborne, in the County of Southampton* (1789), was a triumph of the local and parochial, couched with insight into the universal. Like Jane Austen, White lived through stirring times but mostly kept

quiet about them. His observations on plants, birds, bats and his pet tortoise Timothy (who recognized the hand that fed him) continue to entertain, amuse and instruct.

The year 1789 was a good one for the production of scientists named Bérard. Jacques Étienne Bérard, who died in 1869 (and so achieved a 2.0 and a 1.2), is best known for his work with François Delaroche on the specific heats of gases at constant pressure. He later won a prize for discovering the effects of various gases on delaying the ripening of fruit. Joseph Frédéric Bérard, a physician, worked more in the classificatory and clarifying line. He began his short career with a *Plan d'une médecine naturelle*, and dedicated the rest of it to purging medicine of the misuse of words.

The greatest of French scientists born in 1789 was Augustin-Louis Cauchy, who gave the first comprehensive theory of complex numbers. His numbers were not his only complex. A little schizophrenia appears in his carrying to his first job, as engineer on the Napoleonic naval base at Cherbourg, Laplace's *Mécanique céleste* ("Sire, I have no use for that hypothesis") and Thomas à Kempis's *Imitatio christi*. In 1830 he chose religion over mathematics, refused the oath to the July monarchy, left his professorships and fled to the Jesuits in Fribourg. He returned to Paris in 1839 (1.5) and continued to pour out papers on all mathematical subjects; in less than 20 years no fewer than 589 (14.0) were published in the *Comptes Rendus* of the Paris Académie des Sciences.

1739 (2.5 centenary)

The year 1739 saw the foundation of the Swedish Academy of Sciences. It resembled the British, rather than the Continental, form of such societies in receiving from its king little more than the right to call itself royal. It emphasized practical arts as strongly as basic science. Among its earliest distinguished members were Anders Celsius, of the fixed-point centigrade thermometer (0° = temperature of boiling water and 100° = that of melting ice, a scheme that was later turned on its head); Carl Linnaeus, of the rigid-sex system of botanical classification; Märten Triewald, who had learned his science and technology in England; and Emmanuel Swedenborg, who got his from angels. The Academy is now best known as the awardee of the Nobel prizes in physics and chemistry, and of the prize in Nobel's name in economics. In discharging this onerous responsibility, it can draw on 250 years' experience in balancing the pure and the applied, and on advice ranging from the angelic to the all-too-human.

The applied scientist might also celebrate 1739 for the publication of the second and last volume of Bernard Forest de Bélidor's *Architecture hydraulique*. This

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

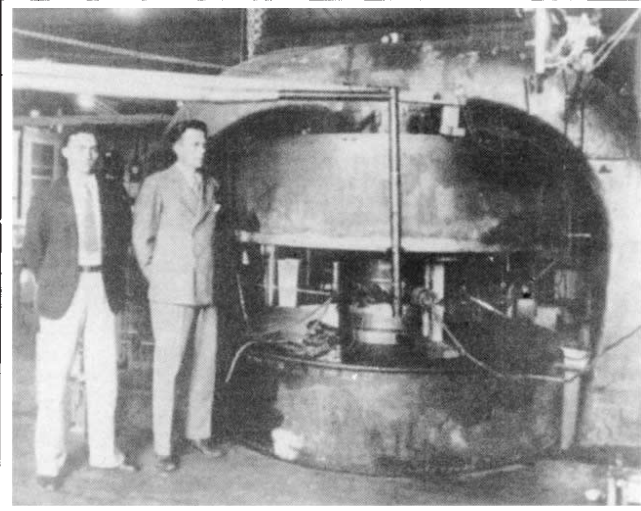
*Lavoisier — prime mover in one revolution,
victim of another.*

work rationalized civil engineering, in which shipbuilding and waterways then played a prominent part, with exemplary constructions and elementary arithmetic. Bélidor had previously written *La science des ingénieurs* — 1729 (2.6) — which deals almost exclusively with fortifications. It indicates an advance in professionalization, and also perhaps in civilization, that 'engineering' is no longer synonymous with 'military architecture'.

1689 (3.0 centenary)

One of the greatest of all describers of disease, Thomas Sydenham (the 'English Hippocrates'), died on 29 December 1689. His happiest days had been spent as a young cavalry officer on the side of Parliament during the English Civil War, before settling down to become the leading London physician of his day. He described his own gout with especial feeling, but his portraits of smallpox, measles, hysteria and fevers mark him as a kind of Gilbert White of medicine. He brought an element of common sense into therapeutics, popularized the use of Peruvian Bark (which contains quinine) for intermittent fevers (the ague), invented a liquid preparation of opium (laudanum) and praised the value of fresh air and exercise. He had friends (including Robert Boyle and John Locke), but was not a clubbable man; above all, he believed in his own senses, which meant he had little truck with microscopes and other newfangled aids.

Consequently, Sydenham had no appreciation of the work of the Dutch lens grinder, Antoni van Leeuwenhoek, whose activity of 1689 included a study of the eye and the identification of the rods of the retina: seeing how we see. ►



Man and machine — Ernest Orlando Lawrence (right) and the 37-inch cyclotron at the University of California, Berkeley, in 1936. This device was one of several of increasing power that Lawrence developed during the 1930s; his inventiveness was rewarded with the Nobel prize in physics for 1939.

1589 (4.0 centenary)

Armchair explorers should bestir themselves to celebrate the four hundredth anniversary of two books that immensely enlarged the knowledge and misinformation of the inquisitive reader of the late sixteenth century. One, Richard Hakluyt's *The principall navigations, voyages and discoveries of the English nation*, amounted to a handbook of the discoveries of the pushy navigators of Elizabethan England. Hakluyt understood the importance of acquainting his countrymen with the elements of geography; and for all the errors of his maps and reports, his readers knew more about the lie of the land than do most British and American high-school students, who cannot find London on a map or name the northern neighbour of the United States. The second mind-expanding book of 1589, Giambattista della Porta's *Magiae naturalis*, is a hodge-podge of recipes, parlor tricks, gee-whiz experiments and natural-philosophical claptrap. "Whatever was notable", says he, "and to be desired through the world, for curiosities and excellent things, I have abundantly found out"; which is the same principle on which this catalogue of anniversaries has been compiled.

1939 (0.5 centenary)

The fireworks have already started for the celebration of the fiftieth anniversary of the discovery of nuclear fission. As noticed in this column last year, Otto Hahn — born 1879 (1.1) — and Fritz Strassmann made and interpreted the critical radiochemical observations in December 1938. In February 1939, Lise Meitner and Otto Frisch — died 1979 (0.1) — published their idea of the physics at work, and appropriated the biological term 'fission'. Immediately physicists and journalists talked up the possibility of the practical application of atomic energy. Would it be Utopia or Armageddon? Also in February, Niels Bohr published his suggestion that the rare lighter isotope of

uranium, uranium-235, was the source of the fission. It therefore appeared that the exploitation of the atom required the separation of heavy isotopes in bulk. Doubt about the possibility of effecting this separation persisted through the year, encouraging those who feared a bomb and frustrating those who hoped for power.

The Nobel prize in physics given during the bicentennial year of the Royal Swedish Academy of Sciences went to Ernest O. Lawrence for his invention of the cyclotron. By distinguishing the inventor of a machine and a method, rather than an investigator who made discoveries, the prize certified the arrival and the importance of the big, interdisciplinary, academic physics laboratory. The 1948 Nobel laureate for physiology or medicine was a chemist who made his observations in 1939. Paul Müller — born 1899 (0.9) — culminated his search for the perfect insecticide with the discovery that 4,4'-dichloro-diphenyl-trichloroethane satisfies many of the criteria. The patent on DDT, as he called it, was secured in 1940 and by the end of the Second World War the miracle chemical was available to help with programmes of insect control and disease eradication. The problems did not emerge until later. Other harbingers of the post-war world also appeared: in April 1939, RCA put its first television sets on the market; in August the German aircraft maker Heinkel flew the first jet aeroplane, too late to be operational for the Battle of Britain; and in December, a company incorporated to exploit Edwin H. Armstrong's patents for frequency-modulated radio, in good time to fight with RCA over the control of frequency bands.

It was already abundantly clear by 1939 that the war to end all wars had failed. Among those who did not live to see the horror of the coming years were Sigmund Freud, Harvey Cushing, Karl Friedrich von Auwers, George Barger, Frank Dyson, Søren Sørensen, Floyd Richtmyer and E. B. Wilson.

1914 and 1964 (0.5 ± 0.25 centenaries)

The celebrated experiments of James Franck — died 1964 (0.25) — and Heinrich Hertz, in which they measured the excitation potential of mercury vapour, took place in 1914. The fact that they at first misinterpreted their results as ionization potentials and hence as fatal to Bohr's new theory of the atom is less celebrated. In 1926 they received the Nobel prize "for their discovery of the laws governing the impact of an electron upon an atom", that is, for confirming Bohr's ideas. The detection of the omega minus meson in 1964 represents an entirely different sort of discovery, one driven by a theoretical prediction, in this case the system of quarks which won the Nobel prize for 1969 (0.20) for its inventor, Murray Gell-Mann — born 1929 (0.60). The discovery of CP violation, which also took place in 1964, showed that nature had not yet run out of surprises.

The Japanese contribution to modern science has been substantial. Many of the earliest pioneers who sailed from the Land of the Rising Sun to Europe went to learn the latest bacteriological techniques. Among them were Sahachiro Hata, Hideyo Nogouchi and Shibasaburo Kitasato. The latter worked with both Koch and von Behring and was the founder, in 1914, of the Kitasato Institute in Tokyo. On Christmas Day of the same year, in far-off Rochester, Minnesota, E. C. Kendall crystallized a substance which he called thyroxin. Crude thyroid extract was already used therapeutically but Kendall's achievement was to make possible the synthesis and elucidation of the molecular structure of the active thyroid substance. Kendall shared the Nobel prize in 1950: not for this research, important though it was, but for his work, with Philip Hench, on the isolation and therapeutic uses of 'Compound E' (cortisone), reported by them in 1949 (0.4).

We end, as we began, with the gargantuan. In 1914 the Panama Canal admitted its first ships after a decade of digging. Its completion required the conquest of yellow fever and the removal of 240 million cubic yards of earth. In 1964, the Verazano Narrows Bridge, begun in 1959 (0.3), the longest single-suspension span in the world, opened for traffic. Each of its towers contains steel enough to make 18,000 automobiles, big ones, 1964-vintage; the entire structure weighs almost four times more than the Empire State Building. The atomic bomb exploded that year by the Chinese — their first — would have destroyed both the building and the bridge, had it been tested in Manhattan. □

W. F. Bynum is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BP, UK. J. L. Heilbron is in the Office of History of Science and Technology, University of California, Berkeley, California 94720, USA.

Chemical codes for the control of behaviour in arthropods

Gerd Bicker & Randolph Menzel

Institut für Tierphysiologie-Neurobiologie, Königin-Luise-Str. 28/30, D-1000 Berlin 33, FRG

Neuromodulators and hormones elicit and modify well-defined behaviours. Their mode of action can be studied to advantage in arthropods, where the natural releasing cells and neuronal target circuits are concisely identified. The coordinated actions of biogenic amines and peptides on both central and peripheral neural activity and metabolic processes bias the whole organism to perform a coherent behavioural routine.

THE Greek physician Galen believed that the brain, with its fluid-filled cavities, was mainly a secretory organ that regulated all aspects of health, mood and nervous function. The modern view of the brain stresses that its computational tasks are based on the connections between its constituent nerve cells. But chemistry plays a vital role in expressing this connectivity, the extent of which is only just being appreciated. Although the importance of chemical transmitters for mediating information transfer at the synaptic connections between neurons is well established, in the past few years it has become apparent that this transmission can be altered or modulated by other chemicals, also produced by neurons, known as neuromodulators.

The actions of neuromodulators at the cellular level and their effects on behaviour are known in some detail but it has been difficult to bridge the gap between these two levels. Here we review recent work on crustaceans and insects which is demonstrating how modulators affect the connectivity in neural circuits, tuning and orchestrating their outputs to fit the immediate needs of the animal.

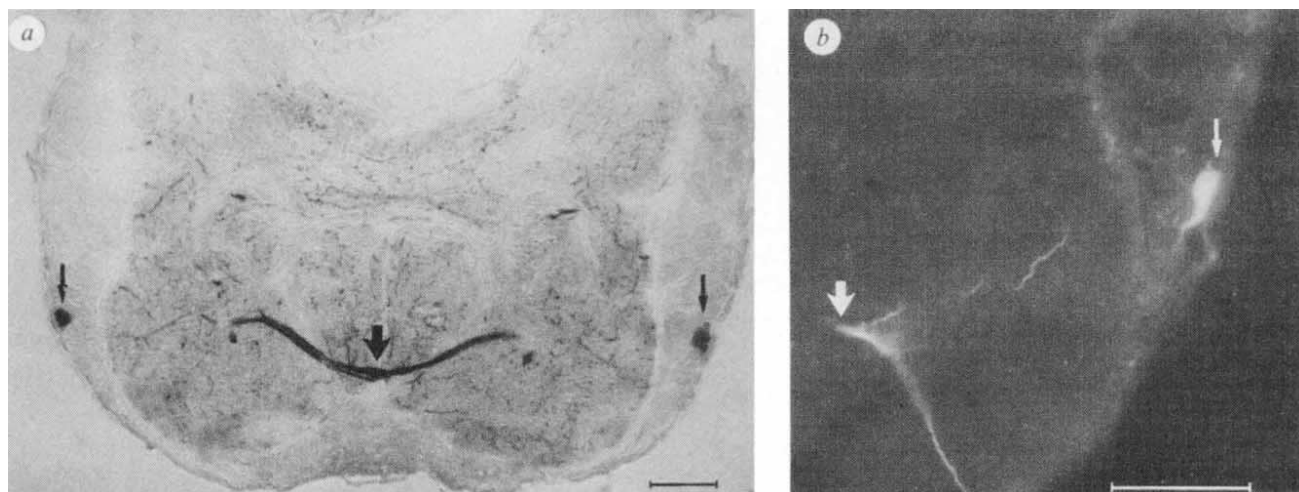
The production of behavioural sequences requires the temporal and spatial coordination of many neural circuits. In addition, there has to be flexibility in the ways in which circuits can be combined because sensory, central and motor neurons can participate in various combinations in many different behaviour patterns. The work in crustaceans and insects is showing that

neuromodulators promote this plasticity by concordantly altering the responsiveness of neurons and muscles, together with changes in metabolism. Through the interplay of several modulators, basic circuitry can be allocated in a variety of ways at different times.

Transmitters, modulators and hormones

The 'classical' transmitters such as acetylcholine, glutamate and GABA directly affect the operation of ion channels and thus rapidly either excite or inhibit the postsynaptic cell. In contrast, the neuromodulators, which include biogenic amines and peptides, act indirectly on the ion channels, modifying their responses to the classical transmitters on a longer time scale. The neurohormones, a third class of chemical messengers, work in a similar way to the modulators, but usually over yet longer times and on targets at a distance from their release sites.

In mammals, neuromodulators and neurohormones are involved in generating specific behavioural states; experimental alteration of neuromodulator levels can evoke specific body movements, postures and motivational states, and alter sensory processing. For example, depletion of the biogenic amine 5-hydroxytryptamine (5-HT, or serotonin) in the brain leads to an increase in arousal and motor activity, evidence that 5-HT normally has a depressing or braking effect on behaviour¹. On the other hand, stimulation of pathways that release another biogenic amine, noradrenaline, increases activity and may



Neurons in the insect nervous system can be identified by a combination of immunohistochemistry and intracellular injection of a fluorescent dye marker. *a*, Serially homologous neurons in the suboesophageal ganglion of the bee that react with antibodies to 5-HT (peroxidase-antiperoxidase reaction). *b*, One of these neurons after intracellular injection with Lucifer yellow. Subsequent processing of this preparation with antibodies to 5-HT confirms that this is one of the neurons shown in *a*. These neurons may be involved in modulating the proboscis-extension reflex, discussed on page 37. Scale bar, 50 μ m. Small arrows: cell bodies; broad arrows: commissural fibres. Photos: Martin Hammer.

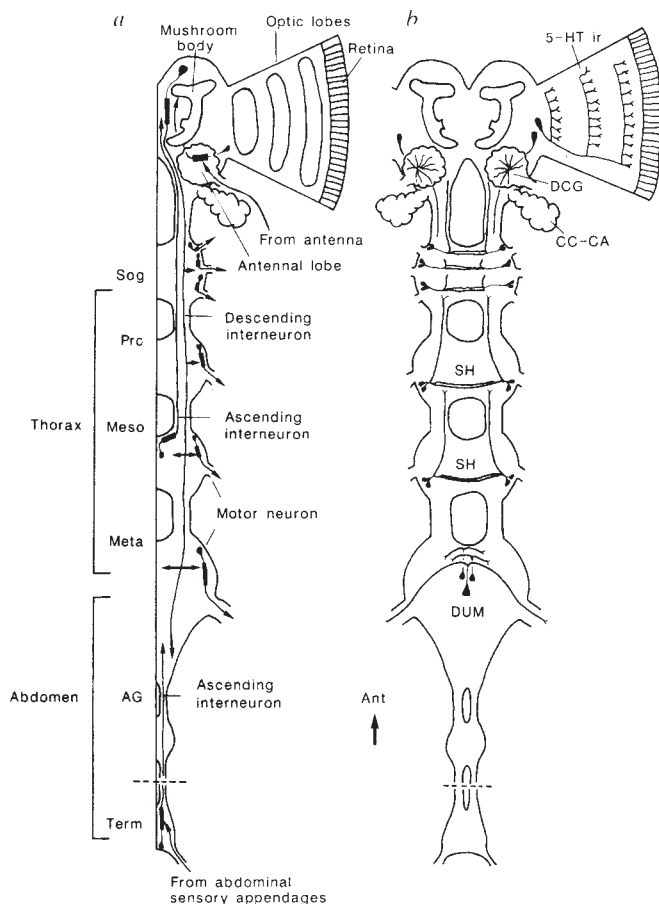


Fig. 1 Schematic drawing showing the organization of the segmented insect nervous system (*a*) and some pathways for information flow (*a* and *b*). The insect nervous system serves as a model of the arthropod nervous system, which consists of an anterior cerebrum (brain) and a ventral chain of segmented ganglia joined by paired connectives. The brain processes sensory information from the retina of the compound eye in the optic lobes and from sensory receptors on the antenna in the antennal lobe. Multimodal integration takes place in more central neuropiles, for example, the mushroom bodies, the output of which is transmitted to the motor centres of the ventral nerve cord by descending interneurons. The brain is linked by paired connectives to the suboesophageal ganglion (Sog) and connects to the hormone-producing neurohaemal organs, the corpora cardiaca (CC) and corpora allata (CA). The suboesophageal ganglion contains the motor circuits for feeding and collaborates with the brain in the initiation, maintenance and coordination of behaviours⁷¹. The ganglia of the ventral nerve cord contain the neural circuitry controlling behaviours such as flight, jumping, respiration, stridulation and oviposition, including the motor neurons that form the outputs to the muscles. These ganglia communicate with the brain through ascending neurons, so that the flow of information in the segmented nervous system is organized in a loop-like fashion⁷¹. Some identified classes of monoaminergic neurons interconnecting several neuropilar compartments, or even segments, are shown schematically in *b*. These include several 5-HT-immunoreactive neurons (5-HT-ir) in the optic lobes⁷², a descending interneuron in the antennal lobe (DCG)⁶⁸ and segmentally homologous neurons (SH)^{68,73}; and a cluster of octopamine-containing unpaired median DUM neurons²⁷⁻³⁰ in each of the thoracic ganglia. Pro: prothoracic ganglion; Meso: mesothoracic ganglion; Meta: metathoracic ganglion; AG: abdominal ganglion; Term: terminal ganglion. Arrows indicate synaptic output areas; thickening of neurites indicates areas of synaptic input. Ant, anterior.

facilitate learning by acting on selective attention or alternatively on stimulus reinforcement². The circuitry in the mammalian brain that regulates these behaviours is, however, so complex that it is difficult to investigate the facilitatory or depressing effects of a neuromodulator in terms of its influence on specific synaptic connections.

Arthropod preparations

In animals such as the crayfish, the lobster and various insects, the circuitry is much more accessible and the cellular mechanisms by which neuromodulators induce behavioural changes can be investigated directly. These animals display some sophisticated behaviours, the generation of which is reasonably well understood at the neuronal level.

The arthropod central nervous system (Fig. 1) is built on a clearly segmented plan and contains many neurons that can reliably be identified from one preparation to another. These are amenable to intracellular microelectrode recording and network analysis, and the interactions between modulatory substances and their target neurons and muscles can be studied on a cell-by-cell basis. In vertebrates the only comparable studies are of very simple behaviours, such as swimming in the lamprey³.

Modulation of rapid-response circuits

In several invertebrate species the experimental activation of single interneurons can initiate complete behavioural acts. One of the best studied examples is the escape behaviour of the crayfish, which begins with a tail flip that propels the animals away from strong tactile stimulation of the tail fan at the end of the abdomen^{4,5}. The inputs from the sensory receptors on the tail fan connect either directly or through a group of sensory interneurons with the lateral giant fibres that mediate the tail-flip response (Fig. 2*a*). When the tail fan is touched, the inputs evoke action potentials in the lateral giant fibres that in turn excite the motor neurons that drive the flexor muscles of the

abdomen. The resulting tail flip pitches the animal away from the source of stimulation⁴.

The threshold for triggering the response varies with the behavioural context: it becomes more difficult to evoke a tail flip when a crayfish is restrained or when it is feeding, and the threshold for triggering action potentials in the lateral giant neurons goes up under these conditions. In the sensory pathways, sensitivity can be altered by repeated tactile stimulation which leads to a gradual decrease in the probability that the reflex will occur. This so-called habituation results from depression of the synaptic transmission between the tactile input neurons and sensory interneurons⁴. Conversely, the probability of eliciting a tail flip can be enhanced or sensitized, if the animal is subjected to strong electric shocks, because the firing threshold in the largest sensory interneuron decreases⁶.

Both habituation and sensitization can be mimicked by applying biogenic amines to the nervous system: 5-HT reproduces the depressive effects of habituation, whereas octopamine mimics the sensitization produced by a traumatic stimulus. Neither amine elicits or affects the performance of the escape behaviours but, rather, they modulate synaptic transmission in the sensory circuits. The excitatory postsynaptic potentials in the lateral giants produced by stimulating sensory neurons are reduced when 5-HT is perfused through the arterial system and enhanced by octopamine perfusion⁷ (Fig. 2*a*); destroying cells that are immunoreactive for 5-HT with a specific neurotoxin also removes the inhibitory effect of 5-HT on transmission at the synapses⁸. From such results, it can be assumed that 5-HT and octopamine have modulatory roles in the escape response.

Once it is initiated by the firing of the giant fibres, the escape response goes automatically to completion, so it seems reasonable that only the synapses in the input circuits should be modifiable. In other behaviours, such as aggressive and submissive postures in the lobster, several levels both centrally and peripherally are modulated simultaneously^{9,10}. The injection of

5-HT into freely moving lobsters evokes a sustained flexion of limbs and abdomen resulting in an aggressive-looking stance, whereas injection of octopamine leads to sustained extension producing a submissive-looking posture^{9,10}. The posture evoked by 5-HT results from the contraction of the postural flexor muscles, whereas octopamine produces contraction of the antagonistic extensors. Application of both amines at the peripheral synapses between motor neurons and muscles, however, induces a more vigorous response in both flexors and extensors, so the antagonistic actions of the two amines must result from modulation of circuits in the CNS (Fig. 2*b*). Electrophysiological recordings show that the spike frequency in flexor motor neurons is increased by 5-HT and decreased by octopamine, and vice versa in the extensor motor neurons. A corresponding antagonism is also found in the excitatory and inhibitory motor neurons that innervate each muscle (unlike vertebrates, arthropods have direct inhibitory innervation of their muscles).

Immunohistochemistry shows over 100 cells reacting with antibodies raised against 5-HT, widely distributed throughout the nervous system; experimental attention has focussed on two pairs of these neurons that have cell bodies in the first abdominal ganglion and fifth thoracic ganglion (Fig. 2*c*). Their morphology¹¹ provides some insight into how the opposing motor patterns could be coordinated by actions at both central and peripheral sites: each has two separate sets of endings, one in the peripheral neurosecretory tissue, where chemicals released can enter the blood, and the other in the neuropiles of the CNS, where integration of information for the activation of the flexor motor programme takes place. The specificity for the action of the amine must be in the CNS, whereas the periphery is simply primed to be more responsive to the central command⁹⁻¹¹.

Modulation of transmission at the neuromuscular junction may be quite widespread. Cockroaches, locusts and crayfish all have some motor neurons that release the pentapeptide proctolin¹²⁻¹⁴, as well as a classical excitatory transmitter, which is probably glutamate¹⁵⁻¹⁷. Proctolin does not depolarize the muscle membrane if applied alone; even though its detailed action is different in each animal, its general effects are to enhance contractions evoked by application of glutamate, or to cause sustained tension instead of a rapid twitch.

Selection of activity patterns

Many rhythmic behaviours result from patterns of activity in central pattern generators (CPGs) which consist of neuronal circuits capable of producing motor output patterns in the absence of sensory feedback. Sensory inputs are nevertheless essential for the activation and shaping of a precisely coordinated motor output. In the lobster, food is ground by three teeth in the anterior part of the stomach, known as the gastric mill, and is then pumped and filtered by the rhythmic constrictions and dilations of the posterior or pyloric end of the stomach. The muscles involved in the movements of the stomach are controlled by a CPG of about 30 neurons located in a small nerve centre termed the stomatogastric ganglion, together with some neurons in the commissural ganglia^{18,19}. The rhythm-generating circuits in the stomatogastric ganglion are capable of producing many different output patterns in response to different transmitters in the input fibres^{19,20} (Fig. 3), allowing a cellular analysis of the interactions between neuromodulators and their target neurons. Using an endoscope and video recordings, rhythmic patterns in the outputs of the isolated stomatogastric ganglion have been related to the movements of the gastric teeth in the intact animals²¹. Injections of proctolin into the blood of an intact animal alters the integration in the CPG circuit in such a way that several modes of chewing and grinding movements are produced in a dose-dependent manner. Neuromodulators such as proctolin may not only influence which output is produced but also modulate the strength of the interaction between the gastric and pyloric burst-generating

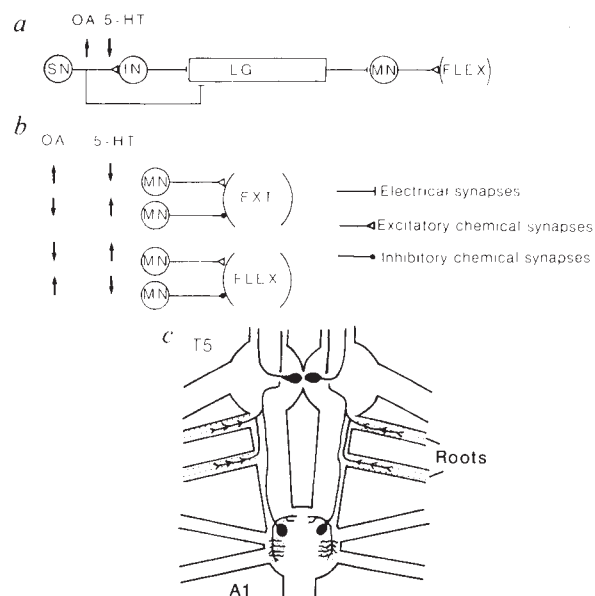


Fig. 2 Neuromodulation of rapid response circuits in crustaceans.

a, In the crayfish tail-flip circuit, transmission at the synapses between the sensory neurons on the tail fan and the sensory interneurons is enhanced by octopamine and depressed by 5-HT and thus alters the threshold of the tail-flip response to tactile stimulation of the tail fan. *b*, Modulation of aggressive and submissive postures in the lobster by octopamine and 5-HT. The two amines have opposing effects on the excitation of the motor neurons of the flexor and extensor muscles. *c*, Schematic drawing of 5-HT immunoreactive cells in the lobster, with somata in the first abdominal ganglion (A1) or the fifth thoracic ganglion (T5). The cells innervate two separate release sites (stippled), a peripheral neurosecretory site in the nerve roots and central sites in the neuropile of the ganglion. SN, sensory neurons; IN, sensory interneurons; LG, lateral giant command fibre; MN, Motorneurons; Ext, Extensor muscles; Flex, fast flexor muscles; Drawings modified after refs 4, 6, 9, 10, 11.

circuits in the stomatogastric ganglion, thus promoting the integrated activity of the stomach when dealing with foods of different qualities.

Orchestration of behavioural units

Fully coordinated sequences of behaviour in several insects can be released by the injection of transmitters or their analogues either into the haemolymph or directly into the nervous system. In general, octopamine has a releasing or enhancing effect that is opposed by 5-HT.

Octopamine not only promotes the expression of a behaviour in insects but also helps to integrate performance at several levels of organization. In restrained locusts, for example, flight motor activity can be initiated by the iontophoretic application of octopamine into the neuropile of the metathoracic ganglion^{22,23}. Octopamine also influences the kinetics of contraction in flight muscles^{24,25}, resulting in an energy saving increase in power output. During long flight sequences, octopamine induces the release of the adipokinetic hormones from stores in the corpora cardiaca, resulting in mobilization of lipid (see Fig. 1)²⁶. Thus, this biogenic amine not only induces flight motor activity in the CNS but also mobilizes the metabolic fuel required for flight.

The main candidates for neurons that release octopamine in the ventral nerve cord of locusts are a group of approximately 90 dorsal unpaired median (DUM) neurons (Fig. 1). Most of these are local interneurons but a few of the larger cells have axons that bifurcate to innervate the musculature on both sides of the body²⁷⁻³⁰. Sombati and Hoyle²² proposed that the release of octopamine from DUM neurons is causally related to the initiation and modulation of behaviour, because iontophoretic

administration in the appropriate ganglia elicits flight or rhythmic stepping movements of the legs, or suppresses the abdominal digging rhythm used by females during egg laying. As electrical stimulation of single DUM cells does not elicit any behaviour^{22,31}, the probability that a behaviour will occur may depend on the pattern of activity over the whole population of DUM cells. The resulting pattern of release in the neuropile could thus determine "which parts of which circuits are most likely to be thrown into (or out of) action"—the so-called orchestration hypothesis²².

Indeed, application of octopamine and its receptor agonists, such as the formamidines, have strong behavioural effects; nocturnal moths become hyperactive during daytime³², the male cabbage looper moth becomes hypersensitive to the olfactory signal that induces sex-pheromone mediated flight, and engorged ticks detach from their hosts³⁴. In the fly and the bee, octopamine or its agonists enhances the reflex response to stimulation of sugar receptors on the feet or antennae with sucrose and causes the animals to feed twice as much as the controls (refs 35, 36 and R.M. and P. Schulz, unpublished results). Because formamidines disturb behavioural co-ordination³⁷⁻³⁸, they may prove to be potent insecticides.

The antagonistic effects of octopamine and 5-HT have been demonstrated by injecting the amines into the haemolymph in several preparations. Octopamine and 5-HT induce opposite postures in males of the cabbage looper moth when they attempt to take off to fly in response to a sex pheromone³³. In honeybees the two biogenic amines may act antagonistically in the visual system. In response to moved striped patterns, bees make direction-specific adjustments of antennal position, a reflex which is enhanced by application of octopamine and reduced by 5-HT in a dose-dependent fashion³⁹. In flight behaviour, different amines may regulate different aspects of the control system: in the hawkmoth *Manduca*, octopamine evokes flight sequences, presumably by enhancing the efficacy of sensory transmission, whereas dopamine seems to have a direct excitatory effect on the interneurons of the flight motor and 5-HT suppresses flight motor output⁴⁰.

Hormonal orchestration of development

The natural release of hormones during the life cycle of an animal provides the most striking examples of the orchestration of behaviour by circulating chemicals over a time span of hours to days. In insects, hormone levels are known to affect reproductive behaviours, such as courtship in crickets⁴¹ or the division of labour in the honeybee hive⁴².

A cellular approach to hormone-mediated behavioural changes during development has been possible in the moth *Manduca*, which undergoes a complete metamorphosis from the larval or caterpillar stage, through the pupa to the adult^{43,44}. The transition through the various stages is regulated by the relative levels of ecdysteroid hormones and juvenile hormones in the haemolymph. A third hormone, eclosion hormone, regulates the final shedding of the pupal cuticle and emergence of the adult, a process termed eclosion. This consists of two phases, the splitting of the cuticle over the thorax by rhythmic movements of the wing bases; and peristaltic waves of the abdomen, which propel the shed cuticle posteriorly. Both phases are centrally programmed, because the appropriate motor patterns can be triggered when the isolated nerve cord is exposed to eclosion hormone.

The nervous system is sensitive to eclosion hormone only shortly before the normal time of hatching. A circadian clock in the brain gates the release of eclosion hormone from neurosecretory centres located in the brain^{43,44}. Eclosion hormone is released as a pulse 2.5 hours before the adult emerges and during this period neurons in the nerve cord become sensitive to the hormone as they start to express two proteins that may be required for the transduction of its signal⁴⁵. The control of the expression of these protein requires the presence of

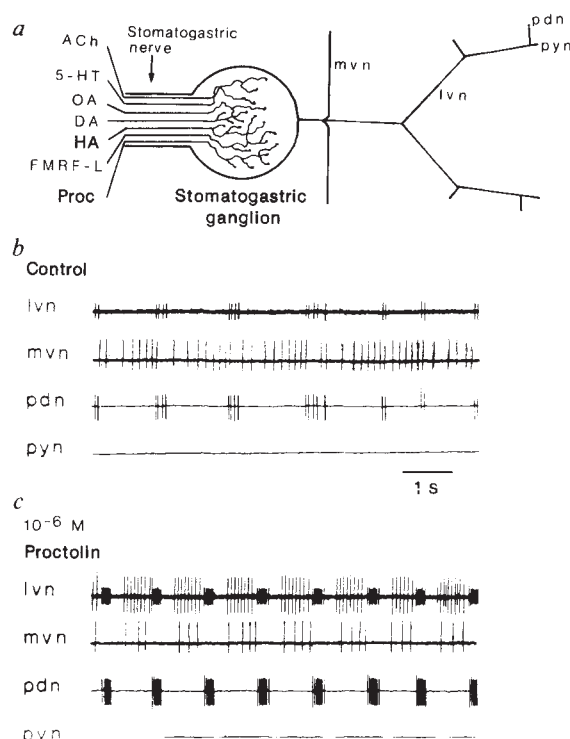


Fig. 3 Neuromodulation of rhythmic motor output of the crustacean stomatogastric nervous system, which control the coordination of the muscles of the stomach. *a*, The stomatogastric nerve carries inputs to the stomatogastric ganglion from other ganglia that use a variety of transmitters. Each transmitter elicits a different motor pattern. Approximately half the cells in the stomatogastric ganglion contribute to the central pattern generator (CPG) for the rhythm of the gastric mill and about the same to that for the pyloric rhythm, both of which can be activated in the isolated ganglion. There is only one interneuron in each of the gastric and pyloric CPGs; the rest of the neurons of the CPGs are motor neurons. The pyloric pattern is an emergent property of the synaptic connections and membrane properties in the network of interneuron and motoneurons and the pacemaker activity of the anterior burster interneuron^{18,19}. *b*, An example of pyloric motor patterns generated by the isolated ganglion, recorded extracellularly from the median ventricular nerve (mvn), lateral ventricular nerve (lvn), pyloric dilator (pdn), and pyloric nerve (pn) shows the modulation of the pyloric rhythm (*c*) when proctolin is added to the bath. It produces an increase in cycle frequency and a change in phase relations in the firing of pyloric motor neurons, which would result in a different pattern of contraction in the pyloric muscles (20). Comparing the effect of proctolin on synaptically isolated neurons with those in the intact network shows that the modulatory effect cannot solely be attributed to direct effects on pyloric neurons (20). Rather, the response of the circuit is shaped by connectivity among neurons that are directly affected and those that are unmodulated ACh: acetylcholine; 5-HT: serotonin; OA: octopamine; DA: dopamine; HA: histamine; FRMF-L: FRMF-like peptide; Proc: proctolin. Modified after refs 19, 20.

ecdysteroid hormone, but also its subsequent removal⁴⁶. Timing is also essential for the sequential activation of the two phases of eclosion behaviour, brought about by a different latency of the action of eclosion hormone on the motor circuits controlling the two phases.

If the pupal cuticle is peeled off before the animal enters the eclosion period, adult behaviours such as walking and the righting reflexes are absent. These appear with a delay of several hours if the eclosion hormone is injected, transforming the prematurely born, uncoordinated animal into an adult with a normal behavioural repertoire⁴³. The changing levels of hormones thus orchestrate profound cellular changes during post-embryonic development that are manifested in stage-specific behaviours. And not only do they affect the responses of neurons

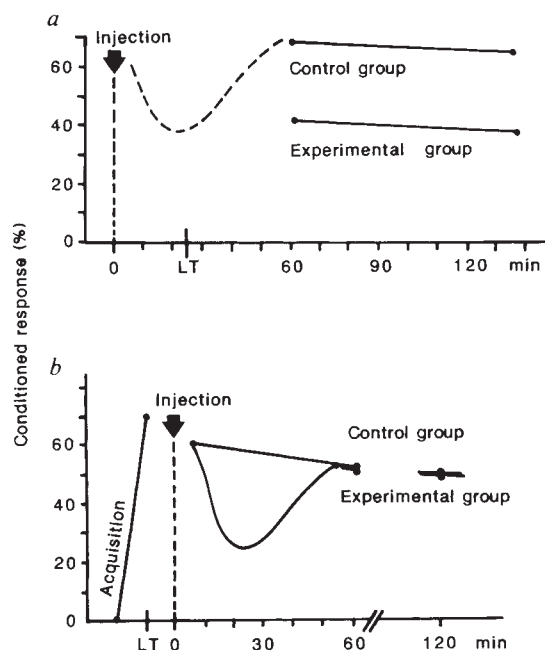


Fig. 4 The experimental procedures used to study the effect of drugs injected into the brain of the honeybee on learning, memory and retrieval from memory. The basic learning procedure is to test a group of bees for their spontaneous responses to an odour and then to condition the non-responding animals once to this odour by pairing the odour with sucrose solution (the learning trial, LT). After conditioning, about 70% of the animals extend the proboscis to the presentation of the odour alone. *a*, Storage test: to see whether a drug interferes with the formation of memory, it is injected before the learning trial. After the drug or a saline control is injected, the animals are conditioned by one trial (LT) at the time when the action of the drug is optimal, as determined in the recall test (see below; dotted curved line indicates that this is a background effect not particularly tested in this experiment). The repeated presentations of odour alone follow when the immediate effect of the drug has vanished. The example given is for an injection of 5-HT (10^{-8} M, 5 nl) into each α lobe of the mushroom bodies (see Fig. 5): the experimental animal shows a persistent reduction in the conditioned response. This proves that 5-HT injected into the α lobe interferes disruptively with the formation of a memory trace. Other drugs like octopamine or noradrenaline may persistently improve the memory when injected before the conditioning trial (see Fig. 5, unpublished results reproduced by kind permission of M. Sugawa). *b*, Retrieval test: the drug is injected into the brain after the learning trial and the animals are tested by repeatedly presenting the odour stimulus alone. The control group (same amount of saline injected) shows a small decline in response, the extinction effect, which is due to repeated testing. The experimental group transiently responds less but the effect of the drug disappears at longer intervals. The example given here is for dopamine (10^{-6} M, 100 nl injected into the protocerebrum), which inhibits conditioned responding maximally at 20–30 min after injection. The animals can be successfully conditioned at a time when the action of dopamine is optimal, showing that the dopamine injection does not interfere with the sensory or motor components or those underlying memory formation⁶¹.

to hormones: they also influence remodelling of neuronal circuits by incorporation of new neurons, reorganization of neuronal arborizations and cell death⁴⁷.

Behavioural plasticity and learning

As neuromodulators confer flexibility on neuronal circuits and trigger intracellular biochemical events that outlast the signal, they have been implicated in learning, memory and the performance of learned tasks in vertebrates^{48–50}, molluscs^{51,52} and fruit flies^{38,53,54}. There is some debate, however, about the site of modulatory action in associative learning.

Classical conditioning is a type of associative learning in

which an animal establishes a connection between a neutral conditioned stimulus (CS), for example, a sound or coloured light and a reinforcing unconditioned stimulus (US) such as food. The repeated temporal pairing of the two stimuli during training elicits a response to presentation of the initially neutral stimulus that is normally caused only by the US. The association of events during classical conditioning requires a locus in the nervous system where the US and CS can converge. An analysis of classical conditioning at the cellular level in the mollusc *Aplysia* has pinpointed an allosteric regulation of an enzyme, membrane-bound adenylyl cyclase, as the most likely point of convergence⁵¹. It is stimulated by the release of a neuromodulator, 5-HT, in the US pathway and its activity is amplified by activity in the CS pathway, which increases intracellular calcium levels.

Some studies in the vertebrate literature, however, consider the US and CS pathways separately from the modulatory functions of biogenic amines, which are assumed to regulate affective states and thus are seen as a third and independent component in the neural substrate of learning^{55,56}. Amines are thought to cause an arousal state in the memory-forming part of the nervous system, so directing 'attention' to the conjugation of stimuli.

The honeybee provides an excellent preparation for addressing this question^{57–62}. Under restrained conditions in the laboratory a memory lasting for days is produced after a single trial in which the CS is an odour, paired with a sucrose reward as the US⁶³. The animal extends its tongue (proboscis), its normal response to sucrose, as a conditioned response (CR) when an initially neutral odour stimulus is presented^{57,64}. This learning process can be manipulated by global or local injections of small quantities of drugs (around 5 nl) directly into the brain. As learning occurs in a single trial that lasts only a few seconds, drugs can be used to separate memory formation from memory retrieval, particularly to examine the unresolved question of whether these are conceptually separate processes. By injecting drugs before the learning trial, storage can be disrupted; injection after the trial disrupts retrieval (Fig. 4).

As an example, consider the action of a global injection of dopamine (10^{-6} M). If it is injected before the learning trial and the animals are tested at a time when the action of dopamine should be optimal (storage test), learning is not impaired but injections after the learning trial (retrieval test, Fig. 4), induces a transient reduction in the conditioned response to the presentation of the CS alone^{58,59,61}. Thus, dopamine interferes selectively with the retrieval process. Because performance in the storage test is unchanged, dopamine has no effect on olfactory coding, motor response or memory storage. In contrast, injections of amines such as 5-HT interfere with the storage process (Fig. 4a) leading to irreversible effects in the conditioned response⁶². As the motor performance during the period when the action of 5-HT is optimal is unaltered, 5-HT must interfere with the storage processes, not with the sensory or motor components of the task.

Local injections of biogenic amines into particular regions of the brain reveal a different pattern of action for each modulator that depends on the area injected and the behavioural components tested⁶² (Fig. 5). For example, in the chemosensory processing area, the antennal lobes, octopamine seems to be involved in stimulus reinforcement. In the higher association areas, the mushroom bodies, octopamine enhances both memory storage and retrieval, whereas noradrenaline enhances only storage and only when injected into the output areas of the mushroom bodies. The facilitatory effects of both octopamine and noradrenaline are antagonized by 5-HT (see Fig. 5 legend for more details).

The most important outcome of these experiments is that learning and retrieval can be improved or depressed by different amines acting at different locations and at different times in the learning process. The experiments also support earlier conclusions^{65,66} that the memory system in the bee brain involves

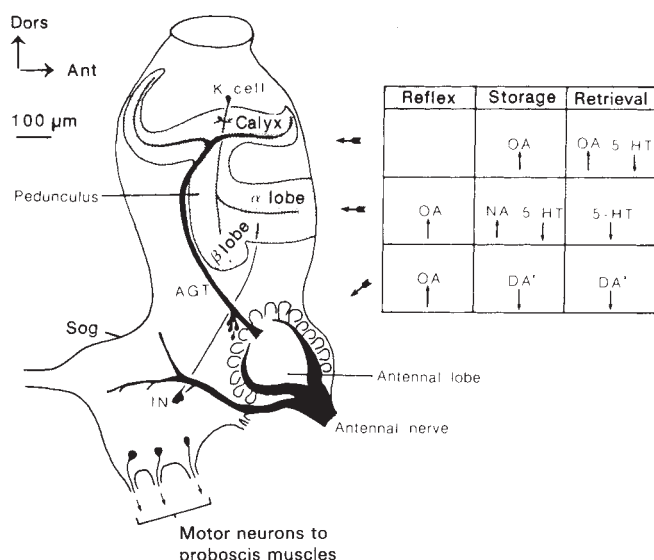


Fig. 5 The effect of local drug injections on the proboscis-extension reflex and on storage and recall of the conditioned response during olfactory learning in honeybees is summarized on a schematic side view of the main pathways in the brain and subesophageal ganglion (Sog) mediating the proboscis extension reflex. Chemosensory afferents from the antenna terminate in the glomeruli of the antennal lobe. Relay neurons transmit olfactory information through the antenno-glomerular tract (AGT) into the calyx of the mushroom body. The mushroom body (stippled) contains approximately 170,000 intrinsic Kenyon cells (K cell), which project from the calyx to the two output regions, the α - and β -lobes⁷⁴. Monoaminergic interneurons (IN) connect the subesophageal ganglion with neuropiles around the α - and β -lobes. The subesophageal ganglion contains the motor circuits and motor neurons responsible for the extension of the proboscis. The table summarizes the action of locally injected amines on the three components of behaviour—proboscis extension reflex, storage and recall. In the table, the arrows pointing upwards indicate facilitation of the reflex or conditioned response, those pointing downwards indicate a reduction. Bold arrows indicate the injection sites. Octopamine (OA) mimics the sensitizing action of the US (sucrose stimulus) and stimulates feeding behaviour when injected into the antennal lobe, suggesting octopamine has a role in reinforcement. If OA is injected into the pedunculus of the mushroom bodies close to the input synapses of the chemosensory relay neurons in the calyx, no changes are found in the reflex performance but both components of learning, memory formation and recall are enhanced. OA has no effect when injected into the output regions of the mushroom body, the α -lobes, whereas noradrenaline (NA) injections into this region enhance memory formation, particularly in animals with lower learning rates. 5-HT on the other hand antagonizes the facilitatory actions of OA and NA in the mushroom body but does not affect the non-associative modulation of the reflex. The substances (2.5 nl , 10^{-6} – 10^{-8} M) were injected into either both antennal lobes, or both α -lobes or both pedunculi (Pd) of the mushroom bodies (the volume of the brains is $\sim 1 \mu\text{l}$). The drug effects were compared with control animals which were injected at the same sites with the same volumes of saline. Several hundred animals were tested^{58,62}. Data for DA* from 60. Dors, Dorsal; Ant, anterior (orientations refer to the position of the brain in the head in life).

several regions, with non-associative modulation of the feeding reflex taking place in the antennal lobes, whereas the mushroom bodies, an area thought in other insect species to be involved in the coordination of behaviour patterns⁶⁷, are necessary for the long-term formation and read out of associative memory. Identified neurons immunoreactive for 5-HT that link these areas have been impaled with dye-filled electrodes⁶⁸, so it should be possible to analyse the learning capabilities of an insect brain at the cellular level.

The distinction between storage and retrieval mechanisms implies that the location of the memory trace is not the direct

sensory pathway between the chemosensory afferents of the antennae and the motor circuit for the proboscis response in the subesophageal ganglion (Fig. 5), because this pathway remains functional when the conditioned response is blocked pharmacologically. Anatomical studies reveal prominent fibre tracts that leave the direct sensory pathway and enter the mushroom bodies (Fig. 5). Storage of the memory trace in a pathway parallel to the main sensory-motor path contrasts the memory system of the bee with the model developed for associative and non-associative plasticity in molluscs, where activity-dependent modulation has been found at the synapses between sensory and motor neurons⁵¹.

Chemical codes and behaviour

Both vertebrate and invertebrate nervous systems contain neurons with widespread branching patterns that seem suitable for coordinating neural activity in a large population of follower cells by a rather diffuse release of neuromodulatory substances. The investigation of some of these identified cells and their targets in arthropod nervous systems suggests that neuromodulators orchestrate behaviour by acting simultaneously at many synaptic levels in the nervous system to integrate neuronal activity, and also by more widespread regulation of metabolism. Sensory circuits may be sensitized or depressed, motor circuits can be biased towards favourable states, muscle tone is adjusted to the correct set point and the metabolism fuelling the muscle power is activated. This coordination of neuronal activity in different parts of the organism has also been observed in molluscs⁶⁹ and leeches⁷⁰, where identified amine-containing neurons integrate several appetitive and consummatory components of feeding behaviour, such as food arousal, swimming toward food sources, bite frequency and meal size.

Experimental manipulation using neuromodulators or their analogues to mimic the natural release of neuromodulators can evoke coherent behaviour. The chemical code for regulating behaviour arises from the patterns of release of the various modulatory substances. These in turn are decoded by the distribution in the target tissue of receptors specific for each type of modulator. As the receptors are linked to the intracellular biochemical signalling pathways, the neuromodulators can modify the excitability of nerve cells on a longer time scale than that of conventional rapid synaptic transmission, leading to longer lasting functional changes in neural circuitry that serve processes such as arousal, selective attention, reinforcement and retention of memory.

Although the behavioural responses to neuromodulators are becoming increasingly better understood, the sensory pathways for activating neuromodulatory neurons are an enigma. A great deal of circuit chasing remains to be done to establish the connectivity between these neurons and external sense organs. Electrophysiological recordings from neuromodulatory neurons of motor circuits^{11,22,31} has so far provided little evidence for a direct link to the mainstream of rapid sensory information processing. But it is conceivable that some neuromodulatory neurons may form, remote from sensory input, a central system for integrating the activity states of many neural circuits. Thus, a release of neuromodulators may mediate the translation of the activity pattern of many neurons into the activity in biochemical pathways in single target cells.

We thank B. Michelsen, P. Rüffer and M. Sugawa for permission to cite their unpublished results, S. Schaare for photographic assistance, M. Whitfield for help with the manuscript and H. J. Pflüger for helpful comments and discussion.

1. Mason, S. T. *Catecholamines and Behaviour*. (Cambridge University Press, London, 1984).
2. Hall, R. D. *Neurosci. Res. Prog. Bull.* **15**, 133–144 (1977).
3. Grillner, S. *Science* **228**, 143–148 (1985).
4. Wine, J. J. & Krasne, F. B. in *The Biology of Crustaceans* Vol. 4 (eds Sandeman, D. C. & Atwood, H. L.) 241–292 (Academic, New York, 1982).
5. Krasne, F. B. & Wine, J. J. *J. exp. Biol.* **63**, 433–450 (1975).
6. Krasne, F. B. & Glanzman, D. L. *J. Neurosci.* **6**, 1013–1020 (1986).
7. Glanzman, D. L. & Krasne, F. B. *J. Neurosci.* **3**, 2263–2269 (1983).

8. Glanzman, D. L. & Krasne, F. B. *J. Neurosci.* **6**, 1560-1569 (1986).
9. Kravitz, E. A. *et al.* in *Model Neural Networks and Behaviour* (ed. Selverston, A. I.) 339-360 (Plenum, New York, 1985).
10. Kravitz, E. A. in *Fast and Slow Chemical Signalling in the Nervous System* (eds Iversen, L. L. & Goodman, E.) 244-259 (Oxford Science, Oxford, 1986).
11. Beltz, B. S. & Kravitz, E. A. *J. Neurosci.* **7**, 533-546 (1987).
12. O'Shea, M. & Bishop, C. A. *J. Neurosci.* **2**, 1241-1251 (1982).
13. O'Shea, M. in *Model Neural Networks and Behaviour* (ed. Selverston, A. I.) 401-413 (Plenum, New York, 1985).
14. Bishop, C. A., Wine, J. J., Nagy, F. & O'Shea, M. R. *J. Neurosci.* **7**, 1769-1779 (1987).
15. Otsuka, M., Kravitz, E. A. & Potter, D. D. *J. Neurophysiol.* **30**, 725-752 (1967).
16. Usherwood, D. N. R. & Cull-Candy, S. G. in *Insect muscle* (ed. Usherwood, D. N. R.) 207-280 (Academic, London, 1975).
17. Bicker, G., Schäfer, S., Ottersen, O. P. & Storm-Mathisen, J. *J. Neurosci.* **8**, 2108-2122 (1988).
18. Selverston, A. I., Miller, J. P. & Wadepuhl, M. in *Symp. Soc. exp. Biol.* (ed. Roberts, A. & Roberts, B. L.) 55-87 (Cambridge University Press, Cambridge, 1983).
19. Marder, E. E., Hooper, S. L. & Eisen, J. S. in *Synaptic Function* (eds Edelman, G. M., Gall, W. E. & Cowan, W. M.) 305-327 (Wiley, New York, 1987).
20. Hooper, S. L. & Marder, E. *J. Neurosci.* **7**, 2097-2112 (1987).
21. Heinzel, H. G. *J. Neurophysiol.* **59**, 551-565 (1988).
22. Sombati, S. & Hoyle, G. *J. Neurobiol.* **15**, 481-506 (1984).
23. Stevenson, P. A. & Kutsch, W. *Naturwissenschaften* **73**, 741-743 (1986).
24. Whim, M. D. & Evans, P. D. *J. exp. Biol.* **134**, 247-266 (1988).
25. Malamud, J. G., Mizisin, A. P. & Josephson, R. K. *J. comp. Physiol. A* **162**, 827-835 (1988).
26. Orchard, I. *J. Insect Physiol.* **33**, 451-463 (1987).
27. Hoyle, G. *J. exp. Zool.* **193**, 433-439 (1975).
28. Evans, P. D. & O'Shea, M. *J. exp. Biol.* **73**, 235-260 (1978).
29. Konings, P. N. M. *et al.* *Cell Tissue Res.* **251**, 371-379 (1988).
30. Goodman, C. S. & Spitzer, N. C. *Nature* **280**, 208-214 (1979).
31. Hoyle, G. & Dagan, D. *J. Neurobiol.* **9**, 59-79 (1978).
32. Shimizu, T. & Fukami, J. I. *Int. Pest Control* **23**, 166-175 (1981).
33. Linn, C. E. & Roelofs, W. L. *Archs Insect. Biochem. Physiol.* **3**, 161-171 (1986).
34. Stone, B. F. *et al.* *Pesticide Biochem. Physiol.* **4**, 407-416 (1974).
35. Long, T. F. & Murdock, L. L. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4159-4163 (1983).
36. Long, T. F., Edgecomb, R. S. & Murdock, L. L. *Comp. Biochem. Physiol.* **83c**, 201-209 (1986).
37. Keyserlingk, H. v. *Agric. Zool. Rev.* **1**, 197-254 (1986).
38. Dudai, Y., Buxbaum, J., Corfas, G. & Ofarim, M. *J. comp. Physiol. A* **161**, 739-746 (1987).
39. Erber, J. & Kaulen, P. in *New Frontiers in Brain Research* 247L (eds Elsner, N. & Creutzfeld, O.) (Thieme, Stuttgart, 1987).
40. Classen, D. E. & Kammer, A. E. *J. Neurobiol.* **17**, 1-14 (1986).
41. Lohr, W. & Huber, F. *J. Insect Physiol.* **10**, 13-36 (1964).
42. Robinson, G. E. in *Neurobiology and Behavior of Honeybees* (eds Menzel, R. & Mercer, A.) 226-279 (Springer, Berlin, 1987).
43. Truman, J. W. & Weeks, J. C. in *Model Neural Networks and Behavior* (ed. Selverston, A. I.) 381-399 (Plenum, New York, 1985).
44. Tublitz, N. J., Copenhaver, P. T., Taghert, P. H. & Truman, J. W. *Trends Neurosci.* **9**, 359-363 (1986).
45. Morton, D. B. & Truman, J. W. *Nature* **323**, 246-267 (1986).
46. Morton, D. B. & Truman, J. W. *J. Neurosci.* **8**, 1338-1345 (1988).
47. Levine, R. B. *Trends Neurosci.* **9**, 315-319 (1986).
48. Dunn, A. J. *A. Rev. Psychol.* **31**, 343-390 (1980).
49. Hunter, B. *et al.* in *Handbook of Psychopharmacology* vol 8 *Drugs, Transmitters & Behavior* (eds Iversen, L., Iversen, S. & Snyder, S.) 531-577 (Plenum, New York, 1977).
50. Mason, S. T. *Prog. Neurobiol.* **16**, 263-303 (1981).
51. Kandel, E. R. *et al.* in *Synaptic Function* (eds Edelman, G. M., Gall, W. E. & Cowan, W. M.) 471-518 (Wiley, New York, 1987).
52. Crow, T. & Forrester, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7975-7978 (1986).
53. Tempel, B. L., Livingstone, M. S. & Quinn, W. G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3577-3581 (1984).
54. Tully, T. *Trends Neurosci.* **10**, 330-335 (1987).
55. Kety, S. S. in *The Neuroscience Second Study Program* (ed. Schmitt, F. O.) 324-336 (Rockefeller University Press, New York, 1970).
56. McGaugh, J. L. *A. Rev. Physiol.* **34**, 297-323 (1983).
57. Menzel, R. *Naturwissenschaften* **70**, 504-511 (1983).
58. Menzel, R. & Michelsen, B. in *Learning and Memory: Mechanisms of Information Storage in the Nervous System* (ed. Matthies, H.) 273-275 (Pergamon, Oxford, 1986).
59. Mercer, A. R. & Menzel, R. *J. comp. Physiol.* **145**, 363-268 (1982).
60. Macmillan, C. S. & Mercer, A. R. *J. comp. Physiol. A* **160**, 359-366 (1987).
61. Michelsen, D. B. *Comp. Biochem. Physiol. C* (in the press).
62. Menzel, R., Michelsen, B., Rüffer, P. & Sugawa, M. in *Modulation of Synaptic Transmission and Plasticity in Nervous Systems* (eds Hertting, G. & Spatz, H. Ch.) 333-349 (Springer, Berlin, 1988).
63. Bitterman, M. E., Menzel, R., Fietz, A. & Schäfer, S. *J. comp. Physiol. Psychol.* **97**, 107-119 (1983).
64. Kuwabara, M. *J. Fac. Sci. Hokkaido Univ. Series 6, Zoology* **13**, 458-464 (1957).
65. Menzel, R., Erber, J. & Masuhr, Th. in *Experimental Analysis of Insect Behavior* (ed. Barton-Browne, L.) 195-217 (Springer, New York, 1974).
66. Erber, J., Masuhr, Th. & Menzel, R. *Physiol. Entomol.* **5**, 343-358 (1980).
67. Huber, F. in *Invertebrate Nervous Systems* (ed. Wiersma, C. A. G.) 333-351 (University of Chicago Press, Chicago, 1967).
68. Rehder, V., Bicker, G. & Hammer, M. *Cell Tissue Res.* **247**, 59-66 (1987).
69. Kupfermann, I. & Weiss, K. R. in *Serotonin, Neurotransmission and Behavior* (eds Jacobs, B. L. & Gelperin, A.) 225-287 (M.I.T. Press, Cambridge, Massachusetts, 1981).
70. Lent, C. M. & Dickinson, M. H. *J. comp. Physiol. A* **154**, 457-471.
71. Altman, J. S. & Kien, J. in *Arthropod Brain* (ed. Gupta, A. P.) 265-301 (Wiley, New York, London, 1987).
72. Nässel, D. R. *Prog. Neurobiol.* **30**, 1-85 (1988).
73. Tyrer, N. M., Turner, J. D. & Altman, J. S. *J. comp. Neurol.* **227**, 313-330 (1984).
74. Mobbs, P. G. *Phil. Trans. R. Soc. B* **298**, 309-354 (1982).

ARTICLES

A unique geochemical record at the Permian/Triassic boundary

William T. Holser*, Hans-Peter Schönlaub†, Moses Attrep Jr‡, Klaus Boeckelmann§, Peter Klein†, Mordeckai Magaritz||, Charles J. Orth‡, Alois Fenninger¶, Catherine Jenny#, Martin Kralik**, Hermann Mauritsch††, Edwin Pak‡‡, Josef-Michael Schramm§§, Karl Stattegger¶ & Rupert Schmölle††

* Department of Geological Sciences, University of Oregon, Eugene, Oregon 97403, USA

† Geologische Bundesanstalt, Postfach 154, A-1031 Wien, Austria

‡ Isotope and Nuclear Chemistry Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

§ Institut für Geologie und Paläontologie, Technische Universität Berlin, Ernst-Reuter-Platz 1, D-1000 Berlin 10, FRG

|| Isotope Research Department, The Weizmann Institute, Rehovot 76100, Israel

¶ Institut für Geologie und Paläontologie, Karl-Franzens-Universität Graz, Heinrichstrasse 26, A-8010 Graz, Austria

Musée géologique cantonal, UNIL-BFSH-2, CH-1015 Lausanne, Switzerland

** Geotechnisches Institut, BVFA Arsenal, A-1030 Wien, Austria

†† Institut für Geophysik, Montanuniversität Leoben, Franz-Josef-Strasse 18, A-8700 Leoben, Austria

‡‡ Institut für Radiumforschung und Kernphysik, Universität Wien, Boltzmanngasse 3, A-1090 Wien, Austria

§§ Institut für Geowissenschaften, Universität Salzburg, Akademiestrasse 26, A-5020 Salzburg, Austria

A 330-metre core drilled through the marine Permian/Triassic boundary in the Carnic Alps of Austria allows closely correlated studies of geochemistry, petrography and palaeontology across the boundary. The isotope shifts and metal concentrations are extended, multiple and complex, and do not resemble those seen at the Cretaceous/Tertiary boundary. Both the carbon isotope shifts and the chemical events (including an iridium anomaly) may have causes related to a major regression of the sea.

THE Permian/Triassic (P/Tr) boundary is associated with the largest extinction event in Phanerozoic time¹. It also witnessed important changes in the environment², notably indicated by a large drop in $\delta^{13}\text{C}$ of the surface ocean from the relatively

enriched ^{13}C condition that characterized the late Palaeozoic^{3,4}. A sudden decrease in $\delta^{13}\text{C}$ has also been observed in connection with other extinction events: the Precambrian/Cambrian⁵, Ordovician/Silurian⁶, Frasnian/Famennian⁷, Devonian/Car-

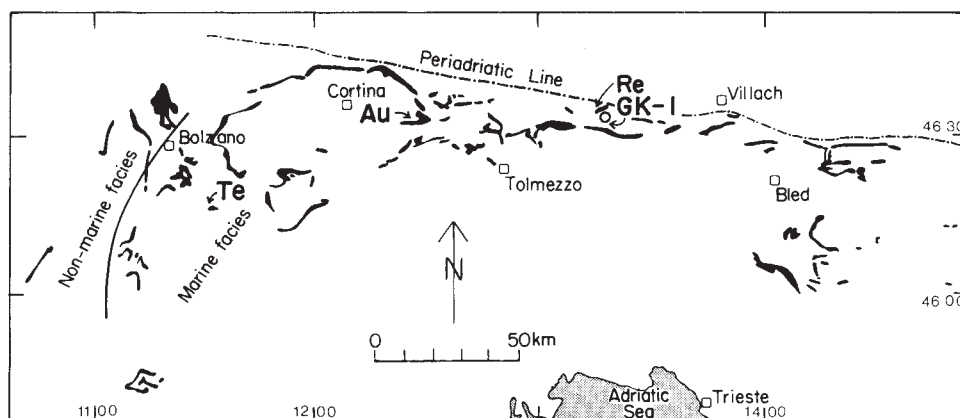


Fig. 1 Location of the Gartnerkofel-1 (GK-1) core hole in relation to the surface distribution of P/Tr rocks in the southern Alps³¹. Also shown are locations of the outcrop sections at the nearby Reppwand (Re), and at Tesero (Te) and Auronzo (Au), Italy²⁸. The GK-1 locality is on a gently dipping marine shelf, 160 km east of the shoreline near Bolzano, Italy.

boniferous⁸, Cenomanian/Turonian⁹ and Cretaceous/Tertiary (K/T)¹⁰. At some of these boundaries, anomalous concentrations of metals, particularly iridium, have been observed. The high concentrations at the K/T boundary¹¹ and the more modest but distinct concentrations in the Late Eocene¹² have been widely ascribed to bolide impacts, whereas concentrations at some of the other extinction events^{7,13-17}, similar to those in the Late Eocene, may have been caused by terrestrial processes (biogenic, oceanic or volcanic).

Suggestions of iridium anomalies associated with the P/Tr boundary have been controversial. Notable iridium anomalies have been reported at exposures of the P/Tr boundary in China¹⁶⁻¹⁹, but analyses of these same sections in other laboratories have failed to find iridium concentrations of more than 100 parts per 10¹² (refs 20-23). Similarly low levels of iridium were found in P/Tr sections in Soviet Armenia²⁴ and in the Dolomite Alps^{25,26}.

Our objective in the current study has been to completely characterize the geochemistry and palaeoenvironment of a representative section of the P/Tr boundary, including carbon isotope and iridium characteristics as well as other measures, and to compare the conditions near the boundary with those in the contiguous sections of the Upper Permian and Lower Triassic.

The Gartnerkofel core

Previous studies of carbon isotope profiles of the P/Tr boundary have indicated that some of the most complete sections across the boundary were situated in the Southern Alps of Austria and Italy²⁷⁻²⁹. To obtain the best possible reference section free of any outcrop alteration, a continuous core was taken specifically for this purpose, on the north slope of the Gartnerkofel near Nassfeld in the Carnic Alps of Austria (Fig. 1). This locality was near the western end of the Tethys Sea, on a carbonate ramp gently inclined to the south-east and 160 km east of the shoreline exposed in the western Dolomite Alps^{30,31}. Hole GK-1 started in the Muschelkalk Conglomerate (Upper Anisian), entered the Werfen Formation (Skythian) at 57 m, the Bellerophon Formation (Dzhufian-Dorashamian) at 231 m, and bottomed in the same formation at 331 m (Fig. 2). Oolites in the lowermost 6 m of the Werfen Formation correlate with the Tesero Horizon in the Italian Dolomites, where the P/Tr boundary, as defined by the last occurrence of Permian fusulinids³²⁻³⁴, lies within this Tesero Horizon. In both GK-1 and the nearby outcrop on the Reppwand cliffs, conodonts were recovered from the lower 40 m of the Werfen Formation. The recognition of *Hindeodus parvus*, a conodont guide fossil for the Early Griesbachian^{35,36}, confirms a very thick and complete representation of lowermost Triassic sediments at this locality. The palaeomagnetic record in the core displays several polarity reversals, as seen in the Permo-Triassic Mixed (Illawarra) Superchron. The reversal nearest to the P/Tr boundary, from positive to negative polarity, lies within the Upper Permian (at

237.2 m); an analogous reversal in the P/Tr section at Shangsi, China, is close to the boundary within a sampling gap of 1.5 m (ref. 37).

The absolute time interval corresponding to the stratigraphic units outlined in Fig. 2 is difficult to assess. Units 1B to 3B—constituting a thickness of ~200 m—correspond to probably not more than one stage each of the Permian (Dorashamian) and Triassic (Griesbachian)³⁸, a time interval of ~5 Myr (ref. 39). This rate of ~40 m Myr⁻¹ is equal to that calculated for the section at Shangsi³⁷, and is within the expected range for carbonate shelf sedimentation⁴⁰.

Closely coordinated sampling of the core allows comparison of a wide variety of data. About 300 billets were cut for thin-section preparation; the remainders of these billets were milled to powder and split for mineralogical, chemical and isotope analyses. Every one of the shaly or marly interbeds, about 50 in number and mostly a few millimetres thick, was similarly sampled and analysed. Samples were spaced in most of the core at a maximum distance of ~1 m, but were as close as 10 cm apart in the vicinity of the boundary. The complete scheme of sampling and analysis, and all the analytical results, are to be published elsewhere (W.T.H. *et al.*, to be submitted to *Jb. geol. Bund.*) Some salient features of the results are given here.

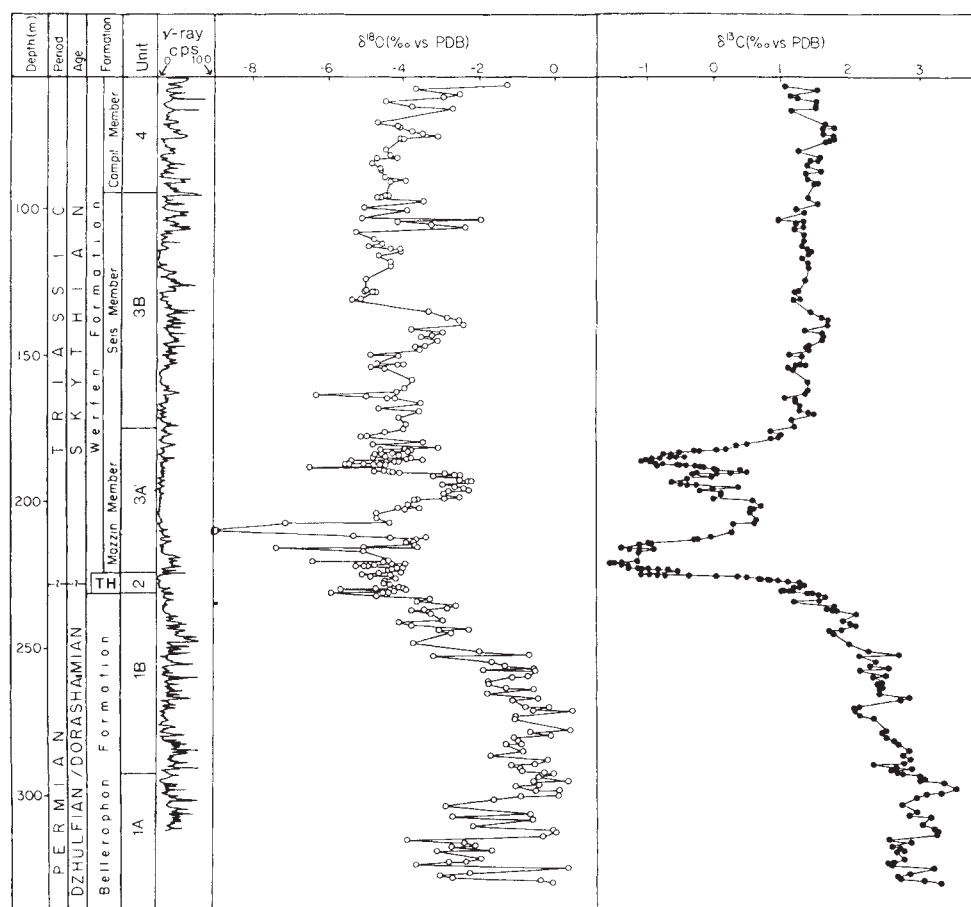
Carbon isotopes

The carbon isotope profile of Figs 2 and 3a suggests further subdivision of the section. Unit 1A, the middle part of the Bellerophon Formation (from the bottom of the core to 293.5 m), has ¹³C-enriched values (mean +3.0 ± 0.3‰). These values are within the range of those reported for Upper Permian rocks, both in other sections in the Alps^{27,28} and throughout Tethys^{29,41}. Unit 1B comprises the upper part of the Bellerophon Formation, where ^δ¹³C values drop gradually from +2.8‰, to +1.5‰ at a depth of 231.3 m, at a rate of 0.02‰ m⁻¹. Comparison of this carbon isotope profile from GK-1 with others to the west at Tesero and Auronzo, Italy (Fig. 1)²⁸, indicates that this segment in GK-1 corresponds to the sharp drop in ^δ¹³C that occurred within a few tens of centimetres in the Italy sections. In Unit 2 (the Tesero Horizon, which includes the P/Tr boundary) ^δ¹³C continues to drop, reaching a value of 0.0‰ at 224.7 m. This segment has the same values of ^δ¹³C and thickness as those of the Tesero Horizon at Tesero^{28,42}. Unit 3A is an interval of low ^δ¹³C in the lower part of the Mazzin Member of the Werfen Formation, extending to 175 m and including three minima, which will be discussed below. This multiple low was not found at any of the other P/Tr sections studied thus far. Units 3B and 4 comprise the upper part of the Werfen Formation, to 57 m, with a constant ^δ¹³C of +1.3 ± 0.2‰.

Oxygen isotopes

Values of ^δ¹⁸O also show some consistent variations in certain intervals (Fig. 2). In Unit 1A the values of ^δ¹⁸O show a bimodal distribution, centred at ^δ¹⁸O = -0.5 and -3.0‰, and the less

Fig. 2 Stratigraphy, geophysical log, and oxygen and carbon isotopes in GK-1. Four lithostratigraphic units are recognized, with subunits based on carbon isotope shifts. Virtually the entire section is fine-grained dolostone, with occasional millimetre-to centimetre interbeds of dolomitic marl or shale (natural γ -ray log >25 counts per second). Unit 1 is the middle and upper part of the Bellerophon Formation, a low-energy, homogeneous, bioturbated biomicroite of a shallow inner shelf, dominated by foraminifers. Unit 2 (231.0–224.5 m), the Tesero Horizon (TH), is transitional to a more shallow and restricted environment: oolites appear in the upper 4 m, with increasing tempestites, intraclasts, ostracodes and worm tubes. Unit 3 is the lower part of the Lower Triassic Werfen Formation, containing strongly bioturbated microsparite, with increasing coquina tempestites. In the interval 95–82 m the section becomes more red, less fossiliferous and richer in terrestrial clastics that characterize Unit 4. At the top of the section a hiatus of the higher parts of the Werfen Formation is followed at 75 m by the Anisian Muschelkalk Conglomerate (not shown). The clastic residues in Unit 1 are mainly illite/muscovite, whose measured crystallinity⁶⁷ attests to only very-low-grade diagenesis (<180 °C). Upwards through Unit 3, detrital quartz and minor chlorite become dominant.



negative values continue in the lower part of Unit 1B. More negative $\delta^{18}\text{O}$ values of about -3.5% occur in the upper part of Unit 1B and persist consistently through Units 2 ($\delta^{18}\text{O} = -4.5 \pm 0.5\%$) and Units 3 and 4 ($\delta^{18}\text{O} = -4.2 \pm 0.8\%$), despite the substantial variations of $\delta^{13}\text{C}$ that characterize those intervals (Fig. 2). The consistent nature of the variations of $\delta^{18}\text{O}$ speaks for a mainly primary record of $\delta^{18}\text{O}$ in this core section. The depletion of $\delta^{18}\text{O}$ across the P/Tr boundary (Units 1 to 2 and 3) may be interpreted as indicative of a temperature rise of perhaps 5°C .

Stable isotopes and chemistry across P/Tr

Detailed profiles of selected stable-isotope and chemical variations for the interval 235–175 m, which crosses the P/Tr boundary, are displayed in Fig. 3. This interval has complex variations of carbon isotopes and chemistry.

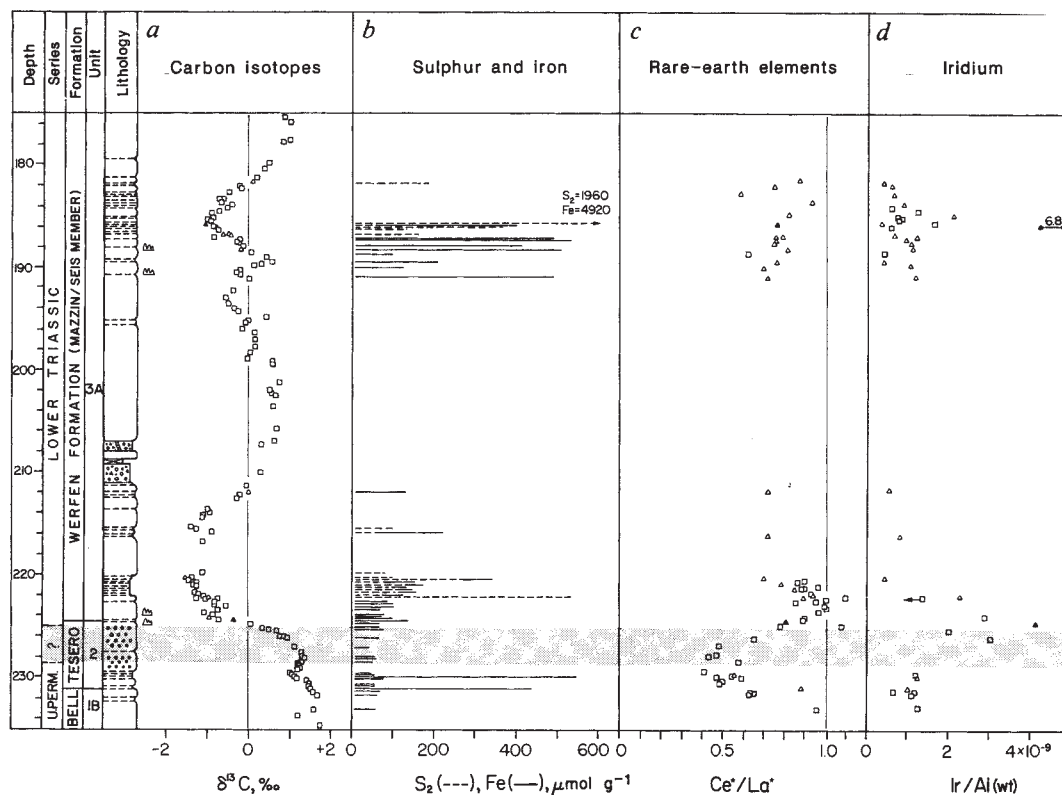
The smooth decrease of $\delta^{13}\text{C}$, which starts in Unit 1B, continues through Unit 2 at an accelerating rate, reaching $0.3\% \text{ m}^{-1}$. This sharp drop in $\delta^{13}\text{C}$ signals the beginning of a zone of low $\delta^{13}\text{C}$ values, which includes three clear minima at 220 (-1.5%), 193 (-0.6%) and 186 m (-1.5%), until a more stable regime with $\delta^{13}\text{C} \approx +1.3\%$ is re-established above 173 m (Fig. 2).

This ^{13}C -depleted zone is also characterized by chemical changes. The first and third carbon isotope minima have high concentrations of pyrite ($>500 \mu\text{mol g}^{-1} \text{S}_2$), as shown by dashed lines in Fig. 3b. Total Fe, shown as solid lines in Fig. 3b, is nearly matched (in $\mu\text{mol g}^{-1}$) by S_2 in these two pyritic intervals, underlain in both cases by oxidized intervals low in S_2 , with most of the iron as goethite. Pyrite has wavy lamellar bedding consisting of globular masses 20–100 μm in diameter, similar to the early diagenetic 'framboidal' pyrite common in modern marine muds and in stratiform sulphide deposits such as the Kupferschiefer, Meggen and Rammelsberg^{43,44}. Each of

the stratigraphically distinct sulphide concentrations in Fig. 3b has a negative, uniform but distinct isotope ratio: for example, near 186 m, $\delta^{34}\text{S} = -19.3 \pm 0.2\%$ and at 221–222 m, $\delta^{34}\text{S} = -25.7 \pm 1.1\%$. These data indicate that the pyrite is syngenetic-diagenetic, and that each bed was deposited in its own distinct reducing environment. This conclusion is supported by rare-earth-element chemistry (Fig. 3c). Ratios of Ce^*/La^* have their highest values of ~ 1.0 in the intervals rich in pyrite, indicating anoxic conditions in the sedimentary regime^{45,46}. Immediately below, in the interval 231–225 m, $\text{Ce}^*/\text{La}^* = 0.5 \pm 0.1$, as would be expected under the highly aerated conditions in which oolites are generated. The rise in Ce^*/La^* begins immediately above the oolite beds, 2 m below the beds in which considerable pyrite is preserved. Above the sulphide zone, Ce^*/La^* ratios (in shaly samples) decrease to intermediate values of ~ 0.7 .

The iridium data also show two peaks (normalized to Al in Fig. 3d) that are substantially above background. The first (165 parts per 10^{12} Ir ($\text{Ir}/\text{Al} = 4.15 \times 10^{-9}$)), at the top of Unit 2 (224.5 m), occurs while $\delta^{13}\text{C}$ is still declining sharply towards its first minimum and the local conditions are just changing from oxic to anoxic. The second Ir peak (230 parts per 10^{12} ($\text{Ir}/\text{Al} = 6.80 \times 10^{-9}$)) is near the top of Unit 3A (185.6 m), centred on the third minimum of $\delta^{13}\text{C}$ and on a peak of S_2 . The highest iridium concentrations are clearly above the relatively low background of iridium ($10\text{--}30 \pm 5$ parts per 10^{12} in shaly samples) that apparently characterizes this geological period; they persist whether plotted as Ir/Al (Fig. 3d), Ir/Fe or just Ir. They are at least an order of magnitude lower than the iridium spike at the K/T boundary, but are similar to iridium concentrations described near other boundaries^{7,13–15}. A log plot of Co against Ir (Fig. 4) shows that Permian/Triassic samples (including both our data from Austria and three analyses from China²¹) have lower Ir than (1) mean CI chondrite⁴⁷, (2)

Fig. 3 Details of stratigraphy and geochemistry in GK-1, from 235 to 175 m, across the P/Tr boundary (shaded). Levels with *H. parvus* indicated alongside lithology. Squares represent dolostone, triangles represent shale or marl, and samples of the two iridium spikes are shown as solid symbols. *a*, Carbon isotope profile; note the steep but smooth decline of $\delta^{13}\text{C}$ across the P/Tr boundary, and the subsequent three minima in the lowermost Triassic; shale and carbonate samples are consonant. *b*, Total sulphur (nearly all is sulphide; analyses by combustion) and total iron (by instrumental neutron activation analysis (INAA)), expressed in $\mu\text{mol g}^{-1}$ of S_2 and Fe, respectively. In all cases the Fe is sufficient to indicate that all S appears as pyrite, FeS_2 , which displays two peaks. *c*, Rare-earth elements (analysed by INAA), illustrated by Ce^*/La^* , each normalized to the North American shale composite⁶⁸. Note the shift from low to high ratio (oxidizing to reducing conditions) at the upper boundary of the Tesero oolite. *d*, Iridium (analysed by RNAA), here normalized by its ratio to Al (analysed by INAA). The low background value of this ratio, at $\sim 1 \times 10^{-9}$, is maintained consistently through all the shaly layers and even above and below the section shown here. Two anomalies of many times the background level, at 4 and 7×10^{-9} , are seen at just above the P/Tr stratigraphic boundary as $\delta^{13}\text{C}$ passes through zero, and at about 40 m above the P/Tr boundary where $\delta^{13}\text{C}$ drops to its third minimum. Each of the two iridium anomalies rises sharply but drops off more slowly, over 1–2 m.



chondritic spherules from the deep sea⁴⁸, (3) chondritic spherules from the Tunguska event⁴⁹, and (4) peak spikes at the K/T boundary^{50–51}. As indicated in Fig. 4, the ratios of Co/Ir in both background and in the nominal peaks of the GK-1 profile are more similar to those found in deep-sea muds⁵² and manganese nodules^{53,54}. Plots of Ni against Ir and Cr against Ir are similar to plots of Co against Ir.

Three samples from the upper Ir high were also analysed by radiometric neutron activation analysis (RNAA) for Au, and gave Au/Ir = 16, 30 and 56 (weight ratio). These are closer to the ratios found in basalts or kimberlites (summarized by Paul *et al.*⁵⁵) than those in chondrites⁴⁷, K/T boundary clays⁵⁰ or deep-sea manganese nodules⁵³.

The weakest carbon isotope minimum (at 193 m; Fig. 3) is in a carbonate interval with no concentration of clay or pyrite, and has not yet been analysed for trace elements by NAA.

Geochemical measures of detrital provenance are constant, with no trend over the whole length of the core: Th/Al = $0.161 \pm 0.028 \times 10^{-3}$ ($n = 76$) and Ga/Sc = 1.70 ± 0.36 ($n = 25$). The latter ratio⁵⁶, and minor olivine and augite in heavy-mineral concentrates in Unit 3, suggest some input of fresh basic volcanic material, such as may have been derived from the Permian/Skythian Haselgebirge of the Northern Calcareous Alps⁵⁷.

Ten analyses of strontium isotope ratios ($^{87}\text{Sr}/^{86}\text{Sr}$) in carbonate rocks between 238 and 185 m (Units 1B, 2 and 3A), corrected for small amounts of radiogenic rubidium, had a mean value of 0.70742 (with a range of 0.70702–0.70777), similar to latest Permian⁵⁸ and earliest Triassic² rocks of western USA, but with no local evidence of an upward trend.

Modelling chemical changes

Taking into consideration the evidence presented here and that in earlier studies^{2,3,28,29}, we suggest the following model for

changes that occurred in the environment across the P/Tr boundary.

The end of the Permian was characterized by one of the strongest regressions of the sea in the entire Phanerozoic², which explains why part of the boundary beds are missing in most of the marine sections around the world. Moreover, the thickness of the section in GK-1 that yields *H. parvus* (Fig. 3a) suggests that the Nassfeld area has a rather complete representation of 'boundary beds'. Thus the chemistry of GK-1 provides a very detailed record which substantially constrains the geochemical model.

A major shift in the carbon cycle was initiated early in the Dorashamian stage, based on evidence in sections across Tethys²⁹. Thus the drop in $\delta^{13}\text{C}$ that begins in GK-1 (at least 50 m below the P/Tr boundary) indicates a net oxidation of the world reservoirs of organic carbon, as sea level dropped, exposing paralic and shallow shelves to erosion. The $\delta^{13}\text{C}$ profile in GK-1 confirms an acceleration in the rate of organic carbon oxidation as the boundary is approached, as was seen previously at Tesero, Italy²⁸. At both Tesero⁴² and in GK-1 (Fig. 3) the occurrence of oolites suggests that the sea had withdrawn to nearly expose these parts of the Tethys shelf.

A sharp change to anoxic conditions is evident at the top of Unit 2 in the GK-1 core by the rise in Ce^*/La^* and then the appearance of pyrite and high S/C ratios. The inception of anoxia, as $\delta^{13}\text{C}$ was dropping through zero for the first time, was marked by anomalous concentrations of iridium and other metals, and the maximum of the anoxic zone corresponds to the initial minimum in the profile of $\delta^{13}\text{C}$. A similar sequence is repeated during a later minimum of $\delta^{13}\text{C}$ at the top of Unit 3A (185.6 m), which is accompanied by another high concentration of metals and sulphide.

A (single) minimum zone of $\delta^{13}\text{C}$ at the base of the Early

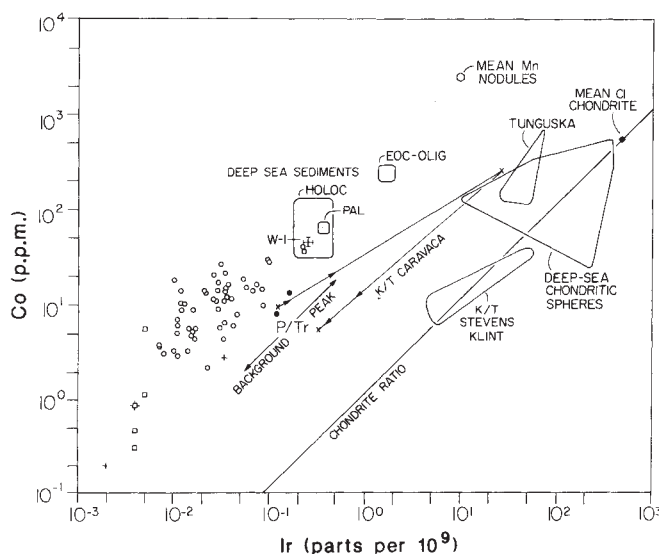


Fig. 4 Plot of cobalt against iridium on a log-log plot. Data from this study (from the Carnic Alps) are indicated as squares (dolomite) or circles (shale or marl); symbols for the two iridium spikes are solid. The data show a general positive correlation, but at a relatively lower Ir value than that for mean CI chondrite⁴⁷ or chondritic spherules^{48,49}. Shown for comparison are two values for the P/Tr boundary beds in Changxing, China²¹ (+), and two examples of the K/T peak (transition through the spike at Caravaca, Spain⁵⁰, marked ×, and the range of peak values at Stevens Klint, Denmark⁵¹). Also shown are approximate ranges of background for Palaeocene and Eocene-Oligocene (ref. 69 and F. T. Kyte and K. T. Wasson, personal communication) and Holocene (Ir: ref. 52; Co: ref. 70) deep-sea sediments, and mean values for manganese nodules (Ir: ref. 53; Co: ref. 54).

Triassic has been verified on a global scale^{3,29}. The chemical changes that we see in GK-1—high iridium, other metals, sulphide and cerium—are certainly an important local signal, although they have not yet been found elsewhere. Whether they are local or worldwide, one must ask why these chemical changes have followed so closely upon the primary, worldwide, longer-term shifts of the carbon cycle. The explanation may be that both the carbon isotope shift and the chemical changes were caused by the same regression of sea level, which first exposed organic carbon to oxidation and finally brought about anoxic conditions to precipitate metals in at least this part of the Tethys Sea. This is consistent with increased erosion of upraised red beds⁵⁹ and volcanics⁶⁰, now exposed west of Bolzano (Fig. 1), bringing metals into solution which were finally reprecipitated on contacting the anoxic waters that had been generated in Tethys. Alternatively, the metals may have been dissolved from the red beds of the underlying Gröden Formation, and precipitated at the oxidation/reduction interface, as has been proposed for the metal-rich (including platinum-group elements) shales of the Kupferschiefer^{61–63}.

Accumulation of metals may have been aided by algae, whose presence is suggested by the lamellar textures seen in this interval. If detailed investigations of other sections show that anoxia and concentrations of metals are localized in this Belleroophon basin, one might then suggest that the cause is an estuarine-style topflow of the fresh waters from the lands to the west during the maximum regression of the Tethys Sea. Alternatively, on a larger scale it could represent a final phase of a deep-sea oxygen-minimum zone that was prevalent for much of the Palaeozoic^{64,65}, brought near the surface during this maximum regression. The prevailing incidence of fresh water eroding continental rocks during this period of regression is evidenced by the high values of ⁸⁷Sr/⁸⁶Sr compared with those of earlier Permian time², but we have not seen a short-term positive spike such as that seen at the K/T boundary⁶⁶.

Unlike other boundaries, for which a single rather sharp drop

in $\delta^{13}\text{C}$ is followed by a relatively rapid recovery to normal $\delta^{13}\text{C}$, we found in GK-1 a thick section (Unit 3A, about 50 m) of low $\delta^{13}\text{C}$, including three separate minima, the two largest of which were characterized by anoxic conditions and enrichment of metals including iridium. Isotope profiles presented previously^{28,29} have indicated that important worldwide changes in the carbon cycle had already begun in Late Permian time and continued smoothly across the stratigraphic P/Tr boundary to culminate in the Early Triassic. The detailed study of GK-1 confirms this extended interval of change, and presents a complex picture of oxidation, reduction and metal deposition constrained within the interval of carbon variations. The evidence presented here, from the most dramatic of all mass extinction events, provides a scenario that is completely different from the bolide impact model proposed for the K/T event, and which is less dependent on iridium anomalies. Certainly, a sharp event or events such as are seen at the K/T boundary are not in evidence here.

The geological coring was funded by the Fonds zur Förderung der Wissenschaftlichen Forschung and the NSF under the Austria-USA Cooperative Science Program. Some of the analytical work was supported by the Geologische Bundesanstalt, NASA and the US Department of Energy, and The Weizmann Institute.

Received 5 August; accepted 14 November 1988.

1. Raup, D. M. & Sepkosky, J. J. Jr *Science* **215**, 1501–1503 (1982).
2. Holser, W. T. & Magaritz, M. *Mod. Geol.* **11**, 155–180 (1987).
3. Holser, W. T., Magaritz, M. & Clark, D. L. *Am. J. Sci.* **286**, 390–402 (1986).
4. Popp, B. N., Anderson, T. F. & Sandberg, P. A. *Bull. geol. Soc. Am.* **97**, 1262–1269 (1986).
5. Magaritz, M., Holser, W. T. & Kirschvink, J. L. *Nature* **320**, 258–259 (1986).
6. Orth, C. J., Gilmore, J. S., Quintana, L. R. & Sheehan, P. M. *Geology* **14**, 433–436 (1986).
7. Playford, P. E., McLaren, D. G., Orth, C. J., Gilmore, J. S. & Goodfellow, W. D. *Science* **226**, 437–439 (1984).
8. Xu, D.-Y. *et al. Nature* **321**, 854–855 (1986).
9. Kauffman, E. G., Elder, W. P. & Pratt, L. M. in *3rd Int. Conf. on Global Bioevents: Abrupt Changes in the Global Biota* 22 (Univ. Colorado, Boulder, 1988).
10. Margolis, S. V., Mount, J. F., Doehe, E., Showers, W. & Ward, P. *Paleoceanography* **2**, 361–377 (1987).
11. Alvarez, W., Alvarez, L. W., Asaro, F. & Michel, H. V. *Geol. Soc. Am. spec. Pap.* **190**, 305–315 (1982).
12. Keller, G. *et al. Meteoritics* **22**, 25–60 (1987).
13. Orth, C. J. *et al. Geology* **16**, 627–630 (1988).
14. Orth, C. J., Quintana, L. R., Gilmore, J. S., Grayson, R. C. Jr & Westergaard, E. H. *Geology* **14**, 986–990 (1986).
15. Orth, C. J. *Geophys. Res. Lett.* **15**, 346–349 (1988).
16. Sun, Y. *et al. 27th Int. geol. Cong. Abstr.* **8**, 309–310 (1984).
17. Xu, D.-Y. *et al. Nature* **314**, 154–156 (1985).
18. Chai, C. *et al. Dizhixue Bao [Acta Geol. Sinica]* **60**, 139–150 (1986).
19. Sun, Y.-Y., Xu, D.-Y., Yan, Z. & Chai, Z. F. *Ber. geol. Bund. (Wien)* **15**, 29–30 (1988).
20. Asaro, F., Alvarez, L. W., Alvarez, W. & Michel, H. V. *Geol. Soc. Am. spec. Pap.* **190**, 517–528 (1982).
21. Clark, D. L., Wang, C.-Y., Orth, C. J. & Gilmore, J. S. *Science* **233**, 984–986 (1986).
22. Zhou L. *Geol. Soc. Am. Abstr. Prog.* **19**, 904 (1987).
23. Boclet, D., Baud, A., Bonté, P., Jéhanno, C. & Rocchia, R. C. *R. Acad. Sci. Paris* **307**, 261–266 (1988).
24. Alekseev, A. S., Barsukova, L. D., Kolesov, G. M., Nazarov, M. A. & Grigoryan, A. G. *Lunar Planet. Sci.* **14**, 7–8 (1983).
25. Oddone, M. & Vannucci, R. in *Field Conference on Permian and Permian-Triassic Boundary in the South Alpine Segment of the Western Tethys, Field Guidebook* 28–29 (Soc. Geol. Italia, Brescia, Italy, 1986).
26. Brandner, R. *Ber. geol. Bund. (Wien)* **15**, 49–56 (1988).
27. Holser, W. T. & Magaritz, M. *Jb. geol. Bund. (Wien)* **128**, 75–82 (1985).
28. Magaritz, M., Bär, R., Baud, A. & Holser, W. T. *Nature* **331**, 337–339 (1988).
29. Baud, A., Magaritz, M. & Holser, W. T. *Geol. Rundschau* (in the press).
30. Buggisch, W. *Carinthia II* **164**, 17–26 (1974).
31. Buggisch, W. *Geol. Rundsch.* **67**, 149–180 (1978).
32. Neri, C. & Passini, M. *Boll. Soc. paleont. Ital.* **23**, 113–117 (1985).
33. Noé, S. *Facies* **16**, 89–142 (1987).
34. Passini, M. *Riv. Ital. Paleont. Stratigr.* **90**, 481–510 (1985).
35. Kozur, H. *Freiberger Forsch. C344*, 85–161 (1978).
36. Sheng, J.-Z. *et al. J. Fac. Sci. Hokkaido Univ. Ser. A* **21**, 133–181 (1984).
37. Heller, F., Lowrie, W., Li H. & Wang J. *Earth planet. Sci. Lett.* **88**, 348–356 (1988).
38. Broglio Loriga, C., Neri, C. & Posenato, R. in *Field Conference on Permian and Permian-Triassic Boundary in the South Alpine Segment of the Western Tethys, Field Guide-Book* 29–34 (Soc. Geol. Ital., Brescia, Italy, 1986).
39. Haq, B. U., Hardenbol, J. & Vail, P. R. *Science* **235**, 1156–1169 (1987).
40. Scholle, P. A., Arthur, M. A. & Edkale, A. A. *Mem. Am. Ass. Petrol. Geol.* **33**, 620–691 (1983).
41. Chen, J., Shao, M., Huo, W. & Yao, Y. *Sci. Geol. Sinica* **1984** (1) 88–93 (1984).
42. Neri, C., Passini, M. & Posenato, R. in *Field Conference on Permian and Permian-Triassic Boundary in the South Alpine Segment of the Western Tethys, Field Guide-Book* 123–128 (Soc. Geol. Ital., Brescia, Italy, 1986).
43. Love, W. G. & Amstutz, G. C. *Fortschr. Miner.* **43**, 273–309 (1966).
44. Ostwald, J. & England, B. M. *Miner. Mag.* **43**, 297–300 (1979).
45. Frey, B. J. *Geochim. cosmochim. Acta* **41**, 361–367 (1977).
46. Wright, J., Schrader, H. & Holser, W. T. *Geochim. cosmochim. Acta* **51**, 631–644 (1987).
47. Wasson, J. T. *Meteorites—Their Record of Early Solar-system History* (Freeman, New York, 1985).

48. Bonté, P., Jéhanno, C., Maurette, M. & Brownlee, D. E. *J. geophys. Res.* **92**, E641-E648 (1987).
49. Ganapathy, R. *Science* **220**, 1158-1160 (1983).
50. Smit, J. & Hertogen, J. *Nature* **285**, 198-200 (1980).
51. Kyte, F. T., Zhou, Z. & Wasson, J. T. *Nature* **288**, 651-656 (1980).
52. Crockett, J. H. & Kuo, H. Y. *Geochim. cosmochim. Acta* **43**, 831-842 (1979).
53. Harriss, R. C., Crockett, J. H. & Stainton, M. *Geochim. cosmochim. Acta* **32**, 1049-1056 (1968).
54. Kelley, V. C. *U.S. geol. Surv. Bull.* **1689A**, (1986).
55. Paul, D. K., Crockett, J. H. & Nixon, P. H. in *Kimberlite Diatremes and Diamonds: Their Geology, Petrology and Geochemistry*, Vol. 1 (eds Boyd, F. R. & Meyer, H. O. A.) 272-279 (Am. geophys. Un., Washington, 1979).
56. Winchester, J. A. & Floyd, P. A. *Chem. Geol.* **20**, 325-343 (1977).
57. Kirchner, E. *Mitt. Österr. geol. Ges.* **71/72**, 385-396 (1980).
58. Brookins, D. G. & Lambert, S. J. *Mat. Res. Soc. Symp. Proc.* **112**, 233-241 (1988).
59. Cassinis, G., Giobbi, E., Neri, C. & Ori, C. G. in *Field Conference on Permian and Permian-Triassic Boundary in the South-Alpine Segment of the Western Tethys*, Field Guide-Book 151-166 (Soc. Geol. Ital., Brescia, Italy, 1986).
60. D'Amico, C. in *Field Conference on Permian and Permian-Triassic Boundary in the South-Alpine Segment of the Western Tethys*, Field Guide-Book 16-22 (Soc. Geol. Ital., Brescia, 1986).
61. Kucha, H. *Tschermak's Miner. Petrogr. Mitt.* **28**, 1-16 (1981).
62. Kucha, H. *Econ. Geol.* **77**, 1578-1591 (1982).
63. Kucha, H. *Trans. Inst. Mining Metall.* **92**, B72-B79 (1983).
64. Goodfellow, W. D. *Chem. Geol. Isotope Geosci. Sec.* **65**, 359-382 (1987).
65. Wilde, P. & Berry, W. B. N. in *Nature and Origin of Cretaceous Carbon Rich Facies* (eds Schlanger, S. O. & Cita, M. B.) 209-224 (Academic, New York, 1982).
66. Macdougall, J. D. *Science* **239**, 485-487 (1988).
67. Kralik, M., Krumm, H. & Schramm, J. M. in *Geodynamics of the Eastern Alps* (eds Flügel, H. W. & Faupl, P.) 164-177 (Deuticke, Vienna, 1987).
68. Gromet, L. P., Dymek, R. F., Haskin, L. A. & Korotev, R. L. *Geochim. cosmochim. Acta* **48**, 2469-2482 (1984).
69. Kyte, F. T. & Wasson, J. T. *Science* **232**, 1225-1229 (1986); *Science* **234**, 1484-1486 (1986).
70. Turekian, K. K. in *Handbook of Geochemistry* (ed. Wedepohl, K. H.) Ch. 27-K, 28-K (Springer, Heidelberg, 1978).

GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose biphosphate carboxylase oligomers in *Escherichia coli*

Pierre Goloubinoff, Anthony A. Gatenby* & George H. Lorimer

Central Research and Development Department, Experimental Station, E.I. DuPont de Nemours and Co., Wilmington, Delaware 19880-0402, USA

Assembly of foreign prokaryotic ribulose biphosphate carboxylases (Rubiscos) in Escherichia coli requires both heat-shock proteins groEL and groES. GroEL is related to a chloroplast protein implicated in Rubisco assembly. Bacteria and chloroplasts therefore have a conserved mechanism that uses auxiliary proteins to assist in the assembly of Rubisco.

THE subunits of some oligomeric proteins lack the inherent ability to correctly assemble into biologically functional molecules¹⁻⁴. For the post-translational assembly of these polypeptides into oligomeric structures a ubiquitous class of conserved proteins, termed 'chaperonins', is thought to be involved¹. This class of proteins includes the *E. coli* groEL gene product¹, as well as homologous proteins in other bacteria^{5,6}, chloroplasts¹, and mitochondria from plants, fungi and animals⁷.

Both groEL and groES proteins are necessary for the morphogenesis of certain bacteriophages⁸⁻¹². They are also involved in DNA replication in *E. coli*^{13,14}. A molecular mechanism accounting for the involvement of the groE proteins in these processes is lacking however. The chloroplast homologue of groEL is called the Rubisco-binding protein because of its demonstrated association with newly synthesized chloroplast-encoded Rubisco large subunits^{15,16}, or with subunits imported into the organelle¹⁷. Both GroEL and Rubisco-binding protein have a similar structure and share extensive amino-acid sequence homology^{1,18}. It has been suggested that in chloroplasts the Rubisco-binding protein is involved in assembly of hexadecameric (L₈S₈) Rubisco¹⁹⁻²¹. Here again, a detailed molecular mechanism is wanting.

Chloroplast Rubisco is comprised of eight large subunits and eight small subunits²². Rubiscos with a similar structure are present in most photosynthetic prokaryotes²³. However, another form of Rubisco, a dimer (L₂) of large subunits, exists in the prokaryote *Rhodospirillum rubrum*²⁴. The dimer from *R. rubrum* is similar in structure to a pair of large subunits of chloroplast Rubisco²⁵⁻²⁷. The structure of the hexadecameric chloroplast enzyme can be generated by tetramerization of *R. rubrum*-like dimers, to which the eight small subunits are polarly appended²⁶⁻²⁸. Both the prokaryotic L₂ and L₈S₈ enzymes can be synthesized and assembled in *E. coli*²⁹⁻³². The demonstrated homology between the *E. coli* groEL protein and the chloroplast Rubisco-binding protein¹, led us to consider the possibility that the assembly of prokaryotic Rubisco in *E. coli* involves the bacterial groEL protein.

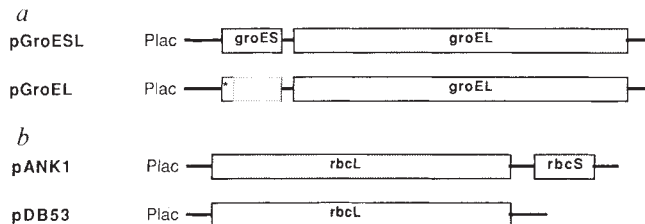


Fig. 1 Plasmid constructions. Two series of compatible plasmids were co-transformed in *E. coli* hosts. *a*, The first series (pGroESL and pGroEL) used plasmid pTG10, which has a p15A replicon, as a parent vector for the groE genes. pGroESL contains the *E. coli* groES and the groEL genes cloned in pTG10. pGroEL contains a frame-shifted (inactivated) groES gene followed by a wild-type groEL gene cloned in pTG10. *Translation frameshift. *b*, The second series of plasmids (pANK1 and pDB53) used pUC18 and pUC8 plasmids respectively, as parent vectors for the Rubisco genes. The pUC plasmids have a pMB1 replicon which is compatible with the p15A replicon. pANK1 (ref. 37) codes for both the Rubisco rbcL and rbcS genes from *A. nidulans*, cloned in pUC18 (ref. 38). pDB53 (ref. 31) codes for only the *A. nidulans* rbcL gene, cloned in pUC8 (ref. 39).

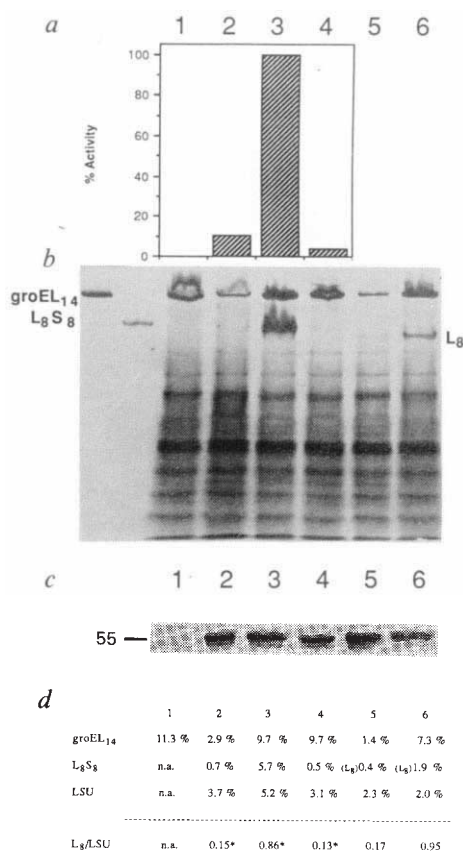
Methods. Plasmid pTG10 was derived from plasmid pACYC184 (ref. 40) which was cut with *Hind*III/*Bam*HI and repaired with Klenow. Into this repaired site was inserted a S1-nuclease-treated 433 base-pair *Hae*II fragment, containing the lac promoter and the multiple cloning site from plasmid pUC8. Plasmid pGroESL was made by cloning the groE operon, as a 2.3 kb *Eco*RI-*Hind*III fragment from plasmid pOF39 (ref. 13), into pTG10. Plasmid pGroEL was derived from a partial *Sac*II digestion of pGroESL, treated with T4 DNA polymerase. The resulting repair of the *Sac*II site located at codon 37 of groES¹³ produced a frame-shift mutation.

Here, we show that assembly of prokaryotic dimeric (L₂), octomeric (L₈), and hexadecameric (L₈S₈) Rubiscos in *E. coli* requires both groES and groEL proteins. Although the groE proteins have been shown to mediate the assembly of L dimers from monomers, we cannot exclude subsequent involvement of these proteins in the formation of octomers and hexadecamers from these dimers.

* To whom correspondence should be addressed.

Fig. 2 *In vivo* assembly of Rubisco in *E. coli*. A *recA56* isolate of strain MC1061 (ref. 41) was co-transformed with the following plasmids: (1) pUC18 and pGroESL; (2) pANK1 and pTG10; (3) pANK1 and pGroESL; (4) pANK1 and pGroEL; (5) pDB53 and pTG10; (6) pDB53 and pGroESL. *a*, Rubisco activity in the cell lysates. *b*, Non-denaturing 7% polyacrylamide gel electrophoresis of lysate proteins, stained with Coomassie Blue. Purified groEL⁴² and spinach L₈S₈ Rubisco⁴³ are shown on the two left lanes. *c*, SDS-polyacrylamide gel electrophoresis of cell lysates, revealed by immunodetection with an antiserum raised against *A. nidulans* Rubisco large subunit. The large subunit appears as a band of relative molecular mass (*M_r*) 55,000. *d*, [³⁵S] counts incorporated in protein complexes from native and SDS-polyacrylamide gels as in *b* and *c*, respectively as a percentage of total TCA counts. n.a., not available; *, predicted L₈ values from L₈S₈ values; LSU, large subunit.

Methods. Cells grown on 2YT medium (16 g Bacto-Tryptone, 10 g Difco yeast extract, 5 g NaCl per litre) under selective conditions, were harvested at A₆₀₀ = 0.4 and resuspended in 100 mM Bicine (pH 8.0), 10 mM MgCl₂, 30 mM NaHCO₃, 15% glycerol, 1 mM dithiothreitol, 15 mg ml⁻¹ phenylmethyl-sulphonyl fluoride, 0.5 mg per ml lysozyme, 0.5 mg per ml deoxyribonuclease, 0.1 mg per ml ribonuclease. Samples were sonicated twice and centrifuged at 6,000 r.p.m. for 1 min at 4 °C. The protein concentration in the supernatant was measured using a Bio-Rad protein assay and adjusted to 5 mg protein per ml. The Rubisco activity was measured as the formation of acid-stable ¹⁴C during a 20-min assay at 25 °C, using 100 µg of protein extract in 0.2 ml of 100 mM Bicine-NaOH (pH 8.0), 10 mM MgCl₂, 1 mM dithiothreitol, 0.4 mM ribulose 1, 5-bisphosphate and 30 mM [¹⁴C]NaHCO₃. (In the presence of both the pGroESL and pANK1 plasmids *E. coli* produced 0.26 nmol of active sites per mg protein as measured by [¹⁴C]CABP binding⁴⁴. 100% activity represents 8.0 nmol of CO₂ fixed per mg protein per min. This corresponds to a turnover number of about 0.5 s⁻¹ and suggests that the enzyme may not have a full complement of small subunits). Non-specific ¹⁴CO₂ fixation was measured in the presence of 100 µM carboxy arabinitol bisphosphate, a specific inhibitor of Rubisco³³. The results were corrected for this background fixation. In a separate experiment, cells were grown for 90 min in 2YT medium supplemented by 100 µCi ml⁻¹ of [³⁵S] methionine, collected and lysed as above. Samples were analysed by non-denaturing PAGE on 7% gels and to SDS-PAGE on 15% gels. Bands containing the native groEL and Rubisco proteins were excised from the non-denaturing gel whereas the band containing the Rubisco large subunit was excised from the SDS gel. Counts of individual gel bands were normalized with values obtained for total cell lysates using precipitation in 10% cold trichloroacetic acid.



GroE proteins stimulate Rubisco assembly

To demonstrate that groE proteins are involved in the assembly of Rubisco in *E. coli*, the level of cellular groE proteins was increased by cloning the *groE* genes on a multicopy plasmid. Two sets of compatible plasmids were constructed. The *E. coli* *groE* genes were cloned on a chloramphenicol resistant plasmid containing a p15A replicon (Fig. 1a), and the *Anacystis nidulans* Rubisco genes (*rbc*) were cloned on an ampicillin-resistant plasmid containing a pMB1 replicon (Fig. 1b). The presence of the plasmid encoding *groES* and *groEL* (pGroESL) in *E. coli* increased the amount of groEL protein from its usual level of 1–2% of cell protein, to about 10% (Fig. 2, b, d).

E. coli cells containing the plasmid encoding the *A. nidulans* *rbcL* and *rbcS* genes (pANK1) produced active Rubisco enzyme that accumulated to a level of 0.7% of cell protein during logarithmic growth (Fig. 2, sample 2). By increasing the cellular concentration of groE proteins with plasmid pGroESL, a substantial increase in Rubisco activity was observed (Fig. 2a, sample 3). This activity was inhibited (>98%) by the addition of carboxy-D-arabinitol 1,5-bisphosphate (CABP), a specific inhibitor of Rubisco³³. An increased quantity of holoenzyme was also apparent on non-denaturing gels (Fig. 2b, sample 3). The identity of the gel band was confirmed by immunoblotting, and was found to correspond to 5.7% of the cell protein, as judged by the incorporation of [³⁵S] methionine (Fig. 2d, sample 3). Except for the groEL and Rubisco oligomers, the pattern of other proteins in Fig. 2b remained similar. Despite the 10-fold variation in active assembled holoenzyme, immunoblotting of SDS-polyacrylamide gels revealed that each of the cell lysates contained comparable quantities of Rubisco large subunit polypeptides (Fig. 2c, samples 2 and 3). The incorporation of [³⁵S]methionine into large subunits did not vary by more than a factor of ~2.5 (Fig. 2d). Therefore, the increased amount of assembled and active Rubisco arising from overproducing the

groE proteins could not be attributed to increased expression of the Rubisco genes, or increased protein stability. We conclude that groE proteins are involved in the post-translational assembly of nascent Rubisco polypeptides into a functional holoenzyme.

On structural grounds, we believe that the final step in assembly of hexadecameric Rubisco is the addition of small subunits to a preformed L₈ octamer, and that small subunits are not involved in the assembly of this octomeric core. Evidence for this is provided by the observations that *A. nidulans* L₈ cores are produced in *E. coli* in the absence of small subunits³¹, and that small subunits can be added to pre-formed cores to yield an active enzyme³⁴. To examine the influence of groE proteins on the assembly of the L₈ cores alone, a plasmid (pDB53) was used which encodes the large subunit of *A. nidulans* Rubisco (Fig. 1b). When large subunit synthesis was directed by plasmid pDB53 in a strain with wild-type levels of groE proteins, a small amount of Rubisco L₈ cores were observed (Fig. 2b, sample 5). In a strain overproducing the groE proteins a significant increase in assembled L₈ cores was observed (Fig. 2b, sample 6). The limited activity of the L₈ cores in the absence of small subunits precluded measurements of enzymic activity in unfractionated lysates. However, titration with [¹⁴C]CABP revealed that 0.14 nmol of sites were formed per mg of cell protein. Using the criteria described earlier, similar levels of Rubisco large subunit polypeptides were found between the various SDS-solubilized cell lysates (Fig. 2c, samples 5 and 6), which thus excluded the influence of the groE proteins on the expression or stability of the large subunit. Therefore, the groE proteins are involved in the conversion of nascent polypeptides into L₈ cores.

Mutations in *groE* prevent Rubisco assembly

If groE proteins promote Rubisco assembly, then mutations in the *groE* locus should confirm the role of the groE proteins,

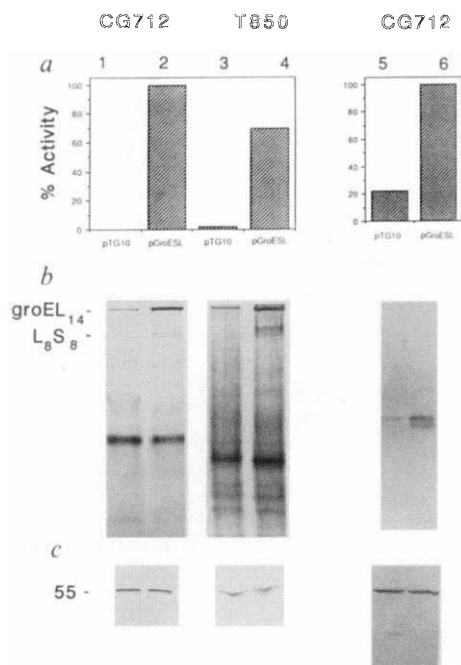


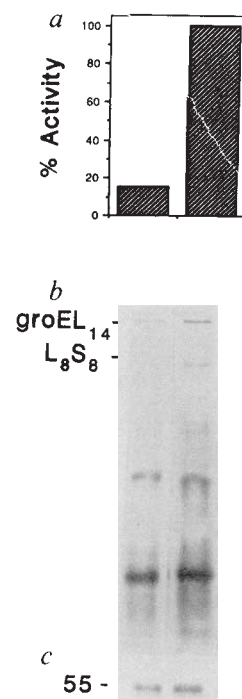
Fig. 3 Suppression by complementation of a deficiency in Rubisco assembly in strains of *E. coli* harbouring defects in the chromosomal *groE* operon. *E. coli groES*⁻ mutant CG712 (ref. 13) or *groEL*⁻ mutant T850 (ref. 45), harbouring *A. nidulans rbcL* and *rbcS* genes on plasmids pANK1 (samples 1–4) were co-transformed with the control pTG10 (samples 1 and 3) or with pGroESL (samples 2 and 4). The *groES*⁻ mutant (CG712) was co-transformed with pRR2119 (ref. 29), which contains the *R. rubrum rbcL* gene, and with either the control plasmid pTG10 (sample 5) or with the pGroESL plasmid (sample 6). **a**, Rubisco activity in cell lysates, *A. nidulans*, samples 1–4; *R. rubrum*, samples 5, 6. 100% activity corresponds for *A. nidulans* Rubisco to 4 nmol CO₂ fixed per mg protein per min, and for *R. rubrum* Rubisco to 10 nmol CO₂ per mg per min. **b**, Non-denaturing PAGE (7% gels) of *E. coli* proteins, stained with Coomassie Blue (lanes 1–4) or revealed by immunodetection (lanes 5,6); **c**, SDS-PAGE (15% gels) of cell lysates, revealed by immunodetection with an antiserum raised against *A. nidulans* Rubisco large subunit M_r 55,000 (lanes 1–4) or against the *R. rubrum* enzyme (lanes 5, 6). The methods for preparing and analysing cell lysates are given in the Fig. 2 legend.

and perhaps identify the relative importance of *groES* and *groEL* in the process. Rubisco assembly was examined in *E. coli* strains with mutations in *groES* (strain CG712) or *groEL* (strain T850), or by using a *groE* overproducing plasmid (pGroEL) which had a frameshift mutation introduced into the *groES* gene (Fig. 1a).

When the *A. nidulans* Rubisco expression plasmid (pANK1) was transformed into CG712 (*groES*⁻), neither Rubisco activity (Fig. 3a, sample 1) nor assembled holoenzyme (Fig. 3b, sample 1) could be detected. The compatible *groES* and *groEL* overproducing plasmid (pGroESL) was introduced into this strain to complement the mutation, which then allowed a significant production of active Rubisco holoenzyme from plasmid pANK1 (Fig. 3a, b, sample 2). A similar result was obtained using the *groEL*⁻ strain T850. Virtually no Rubisco activity (Fig. 3a, sample 3) or assembled holoenzyme (Fig. 3b, sample 3) could be detected when T850 was transformed with the Rubisco expression plasmid (pANK1). When the *groEL* mutation of T850 was complemented by the presence of plasmid pGroESL however, both Rubisco activity (Fig. 3a, sample 4) and assembled holoenzyme (Fig. 3b, sample 4) were observed.

The influence of *groES* on Rubisco assembly was also demonstrated using the dual plasmid expression system. Figure 2 shows that an increased production of the *groE* proteins encoded by pGroESL led to improved assembly of hexadecameric (samples 2 and 3) and octomeric (samples 5 and 6) Rubisco. The *groES*

Fig. 4 Thermo-induction of Rubisco assembly. Wild-type *E. coli* cells (MC1061), harbouring plasmid pANK1, were grown either at 26 °C (left lane) or at 42 °C (right lane) to the cell density of A_{600} 0.4 from a common inoculum, initially grown at 26 °C. **a**, Rubisco activity in cell lysates (100% corresponds to 2.5 nmol of CO₂ fixed per mg protein per min) **b**, Non-denaturing PAGE (7% gel) of cell lysates stained with Coomassie Blue) **c**, SDS-PAGE of cell lysates in 1% SDS, 1% β -mercaptoethanol, revealed by immunodetection with an antiserum raised against *A. nidulans* Rubisco large subunit (M_r 55,000). The methods for preparing and analysing the cell lysates are given in the Fig. 2 legend.



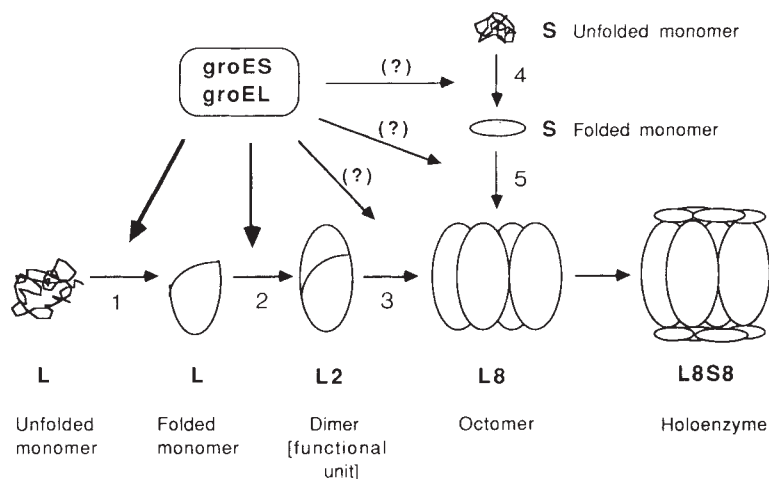
gene of plasmid pGroESL was inactivated by introducing a translational frame-shift to produce plasmid pGroEL (Fig. 1). The presence of the *groES* frame-shift plasmid (pGroEL) and the Rubisco expression plasmid (pANK1) together in *E. coli* did not lead to improved assembly of Rubisco (Fig. 2, sample 4), thus emphasizing the involvement of *groES* in assembly of the hexadecameric enzyme. The *groES* frame-shift did not have a significant polar effect on expression of *groEL* because the plasmid-encoded *groEL* oligomer was present as an abundant protein (Fig. 2b, d, sample 4). Also, synthesis and stability of the Rubisco large subunit were unaffected (Fig. 2c, d, sample 4).

Recent structural studies of the L_8S_8 holoenzyme have revealed that the L_8 core is composed of four *R. rubrum*-like dimers arranged about a fourfold axis, the L_2 dimer being the basic structural and catalytically functional motif (see Fig. 5)^{25–28}. This suggested that assembly of L dimers might be an essential intermediary stage in the formation of the L_8 core, and so the role of the *groE* proteins in this step was examined. The *groES*⁻ *E. coli* mutant (CG712) was transformed with a plasmid (pRR2119) which codes for the *R. rubrum* Rubisco gene. Active, dimeric Rubisco could be detected in extracts from this strain (Fig. 3a, b, sample 5). However, the amount of active dimeric Rubisco was increased almost fivefold when the *groE* expression plasmid (pGroESL) was introduced into the *groES*⁻ strain containing the *R. rubrum* Rubisco plasmid pRR2119 (Fig. 3a, b, sample 6). Again, the level of Rubisco subunits synthesized remained constant despite the different amounts of *groE* proteins produced (Fig. 3c, samples 5 and 6). The *groE* proteins are required for cell viability at 37 °C¹, and the mutant forms of *groES* and *groEL* present in CG712 and T850 would be expected to retain some residual or altered function. This may explain the presence of low levels of active Rubisco detected in *groES*⁻ mutants (Fig. 3a, samples 3 and 5). The enhanced assembly of the L_2 dimer, observed by complementation and overexpression of *groES* and *groEL*, indicated that these proteins functioned at the very least at the stage of dimer formation.

Heat shock improves Rubisco assembly

In *E. coli* the expression of *groES* and *groEL* is stimulated following heat shock³⁵. As assembly of Rubisco is influenced by the *groE* proteins, a heat shock to cells synthesizing Rubisco subunits might result in increased assembly of holoenzyme.

Fig. 5 Proposed assembly pathway of Rubisco hexadecamer. Recognizing that the L_2 dimer is the basic structural motif²⁵, a minimum of five steps must be invoked to convert nascent, unfolded large and small subunits into the hexadecameric L_8S_8 holoenzyme: (1) folding of large subunit monomer; (2) formation of the L_2 dimer; (3) tetramerization of L_2 dimers to give the L_8 core; (4) folding of small subunit monomer and (5) association of small subunits with the L_8 core. The groE proteins could potentially be involved at each step, although evidence³⁶ suggests that step 5 can occur spontaneously without their involvement. The results reported here show that, at the very least, the groE proteins are involved in dimer formation (that is, steps 1 and/or 2). However, the possibility that the groE proteins are additionally involved in steps 3 and 4 cannot be excluded.



A strain (MC1061) with wild-type chromosomal *groES* and *groEL* genes was transformed with the plasmid encoding *A. nidulans* Rubisco genes (pANK1). Cells were grown at 26 °C in liquid culture, and then inoculated into fresh media at either 26 °C or 42 °C. The heat shock increased the synthesis of *groEL* (Fig. 4b), and also resulted in a sixfold increase in Rubisco specific activity (Fig. 4a). Production of the L_8S_8 holoenzyme was also enhanced (Fig. 4b), but there were comparable amounts of large subunits produced at both temperatures (Fig. 4c). When the experiment was repeated with a *groEL*⁻ strain (T850), only background levels of assembled enzyme were obtained (data not shown). The stimulation of Rubisco assembly by heat shock therefore required a functional *groEL* protein. The pleiotropic nature of the heat-shock response requires some caution in interpretation of these experiments. The results are consistent however with the *groE* overproduction experiments and use of mutants described earlier.

Discussion

The experiments described here demonstrate that the *groES* and *groEL* proteins are both obligate accessories for the assembly of prokaryotic Rubiscos in *E. coli*. The assembly of dimeric, octomeric and hexadecameric Rubiscos are all influenced by the *groE* heat-shock proteins. There are several potential sites of action for *groE* proteins in the assembly of Rubisco (Fig. 5). These sites of action might be non-exclusive and could be at the steps of polypeptide-chain folding to form L and S monomers, dimerization of L monomers, tetramerization of L dimers, or adjunction of small subunits to the L octomer. Experi-

mentally, we have shown that the *groE* proteins are required for assembly of the *R. rubrum* L dimer. Dimerization of L could be the basic common step involved in the assembly of the L octomer and the L_8S_8 hexadecamer of *A. nidulans*. Addition of small subunits to the L_8 core does not seem to require *groE* proteins, as this step can be accomplished *in vitro* with no additional protein factors³⁶. We cannot exclude a role for the *groE* proteins in the folding of the small subunits.

Our data shows for Rubisco, and implies for other proteins, that the primary structure of some oligomeric proteins is insufficient to specify their spontaneous assembly into a biologically active form *in vivo*. The dependence of a foreign oligomeric protein on host heat-shock proteins for correct assembly provides experimental evidence for the existence of a conserved mechanism capable of controlling the assembly of a variety of oligomeric proteins. This has also been inferred by structural comparisons of chaperonins¹.

The requirement for *groES* in Rubisco assembly in *E. coli* is of particular interest when considering the biogenesis of organelles in eukaryotes. A *groEL* homologue is present in chloroplasts and mitochondria^{1,7}, and has been implicated in the assembly of plant Rubisco¹⁹⁻²¹. The sequence and functional conservation of *groEL* in protein assembly leads us to predict that a protein with a function analogous to *groES* may be present in organelles to assist the post-translational assembly of some of their proteins.

We acknowledge the help of O. Fayet, C. Georgopoulos, R. W. Hendrix, and S. Gutteridge for providing strains and plasmids, T. H. Lubben for comments on the manuscript and C. Kalbach for technical assistance.

Received 16 September; accepted 7 November 1988.

- Hemmingsen, S. *et al.* *Nature* **333**, 330-334 (1988).
- Pelham, H. R. B. *Cell* **46**, 959-961 (1986).
- Ellis, R. J. *Nature* **328**, 378-379 (1987).
- Pelham, H. R. B. *Nature* **332**, 776-777 (1988).
- Mehra, V. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 7013-7017 (1986).
- Vodkin, M. H. & Williams, J. C. *J. Bact.* **170**, 1227-1234 (1988).
- McMullin, T. W. & Hallberg, R. L. *Molec. cell. Biol.* **8**, 371-380 (1988).
- Georgopoulos, C. P. & Hohn, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 131-135 (1978).
- Zweig, M. & Cummings, D. J. *J. molec. Biol.* **80**, 505-518 (1973).
- Sternberg, N. *J. molec. Biol.* **76**, 25-44 (1973).
- Kochan, J. & Murialdo, H. *Virology* **131**, 100-115 (1983).
- Tilly, K. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 1629-1633 (1981).
- Fayet, O., Louarn, J.-M. & Georgopoulos, C. *Molec. Gen. Genet.* **202**, 435-445 (1986).
- Jenkins, A. J. *et al.* *Molec. Gen. Genet.* **202**, 446-454 (1986).
- Barraclough, R. & Ellis, R. J. *Biochim. biophys. Acta* **608**, 19-31 (1980).
- Roy, H. *et al.* *J. Cell Biol.* **94**, 20-27 (1982).
- Gatenby, A. A. *et al.* *EMBO J.* **7**, 1307-1314 (1988).
- Pushkin, A. V. *et al.* *Biochim. biophys. Acta* **704**, 379-384 (1982).
- Gutteridge, S. & Gatenby, A. A. in *Oxford Surveys of Plant Molecular and Cell Biology* (ed Miflin, B. J.) 95-135 (Oxford University Press, Oxford, 1987).
- Ellis, R. J. & van der Vies, S. M. *Photosynthesis Res.* **16**, 101-115 (1988).
- Roy, H. & Cannon, S. *Trends biochem. Sci.* **13**, 163-165 (1988).
- Miziorko, H. M. & Lorimer, G. H. *A. Rev. Biochem.* **52**, 507-535 (1983).

- Andrews, T. J. & Lorimer, G. H. in *The Biochemistry of Plants* (ed Hatch, M. D.) 131-218 (Academic, New York, 1987).
- Tabita, F. R. & McFadden, B. A. *J. biol. Chem.* **249**, 3459-3464 (1974).
- Schneider, G. *et al.* *EMBO J.* **5**, 3409-3415 (1986).
- Chapman, M. S. *et al.* *Nature* **329**, 354-356 (1988).
- Andersson, I. A. *et al.* *Nature* (in the press).
- Chapman, M. S. *et al.* *Science* **241**, 71-74 (1988).
- Somerville, C. R. & Somerville, S. C. *Molec. Gen. Genet.* **193**, 214-219 (1984).
- Gatenby, A. A., van der Vies, S. M. & Bradley, D. *Nature* **314**, 617-620 (1985).
- van der Vies, S. M., Bradley, D. & Gatenby, A. A. *EMBO J.* **5**, 2439-2444 (1986).
- Tabita, F. R. & Small, C. L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6100-6103 (1985).
- Pierce, J., Tolbert, N. E. & Barker, R. J. *Biochemistry* **19**, 934-942 (1980).
- Andrews, T. J. *J. biol. Chem.* **263**, 12213-12219 (1988).
- Cowing, D. W. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 2679-2683 (1986).
- Andrews, T. J. & Bailment, B. J. *J. biol. Chem.* **258**, 7514-7518 (1983).
- Gutteridge, S., Lorimer, G. H. & Pierce, J. *Pl. Physiol. Biochem.* **26** (in the press).
- Norrander, J., Kempe, T. & Messing, J. *Gene* **26**, 101-106 (1983).
- Vieira, J. & Messing, J. *Gene* **19**, 259-268 (1982).
- Chang, A. C. Y. & Cohen, S. N. *J. Bact.* **134**, 1141-1156 (1978).
- Casadaban, M. & Cohen, S. N. *J. molec. Biol.* **138**, 179-207 (1980).
- Hendrix, R. W. *J. molec. Biol.* **129**, 375-392 (1979).
- Lorimer, G. H., Badger, M. R. & Andrews, T. J. *Biochemistry* **15**, 529-536 (1976).
- Hall, N. P., Pierce, J. & Tolbert, N. E. *Archs. Biochem. Biophys.* **212**, 115-119 (1987).
- Takani, T. & Kakefuda, T. *Nature* **239**, 34-37 (1972).

A second supernova inside Puppis A?

P. F. Winkler*, R. P. Kirshner†, J. P. Hughes‡
& S. R. Heathcote‡

* Department of Physics, Middlebury College, Middlebury, Vermont 05753, USA

† Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA

‡ Cerro Tololo Inter-American Observatory, Casilla 603, La Serena, Chile

Young supernova remnants provide rare glimpses of material from the cores of massive stars, leading to observational tests of stellar evolution and nucleosynthesis theories. Here we present narrow-band images of an unusual swirl structure located near the centre of the Puppis A supernova remnant, which reveal three overlapping filamentary systems of highly diverse composition. One is dominated by emission lines of nitrogen, one by oxygen and one by sulphur. Spectra indicate that these small rings have high velocities and a corresponding kinematic age of less than 800 yr. Heavy-element abundances are extremely high and different in each system, some reminiscent of knots in Cas A, and others of circumstellar matter observed surrounding supernova 1987A. The youth and unusual chemistry lead us to suggest that a second supernova has exploded within the shell of Puppis A, giving rise to the swirl structure. A similar suggestion was made by Elliot *et al.*¹, based only on the region's appearance in $H\alpha$.

The irregular shell structure of Puppis A, about 50 arcmin in diameter, is apparent at both radio² and X-ray³ wavelengths. Its bright optical filaments lie predominantly near the shell's

rim to the north-west and to the east (Fig. 1). Spectra of these bright filaments show emission lines typical of supernova remnants (SNRs), although some are notable for very strong nitrogen lines⁴⁻⁶. Measured radial velocities^{7,8} for these filaments have without exception been modest, $v_r < 300 \text{ km s}^{-1}$. In striking contrast are a recently discovered⁹ set of faint filaments which have strong oxygen lines, almost no Balmer emission, and radial velocities as high as $1,500 \text{ km s}^{-1}$. Proper motions for the latter filaments are consistent with undecelerated expansion from a common origin 3,700 years ago¹⁰. The explosion of a massive star $\sim 4,000$ years ago is consistent with all previous data from Puppis A at X-ray, optical and radio wavelengths.

Our new observations come from a complex region¹ which comprises filamentary systems of three different types, two of which have not been previously observed in Puppis A. This structure is centred near $\alpha = 8^{\text{h}}20^{\text{m}}56^{\text{s}}$, $\delta = -42^{\circ}50'$ (1950), 3 arcmin south-east of the centre from which the fast oxygen-rich filaments are expanding¹⁰, and appears as a swirl of emission ~ 4 arcmin in diameter—less than one tenth of that of Puppis A—on the deep $H\alpha + [\text{N II}]$ plate shown in Fig. 1. No excess emission of the swirl region over its surroundings is seen in radio² or X-ray³ maps of Puppis A.

We used the prime-focus CCD camera with an RCA chip at the CTIO 4-m telescope on 17 February 1988 to obtain narrow-band images of the swirl region at wavelengths of three prominent emission lines plus a continuum image (see Table 1 for details). The four images are shown in Fig. 2. Each of the emission-line images shows a filamentary structure typical of SNRs, yet the images in the different bands are so dissimilar that it would be difficult to recognize them as the same object, were it not for the star field.

Two nights later we used the Ritchey-Chretien spectrograph and 2D-Fruti detector on the CTIO 4-m telescope to obtain a

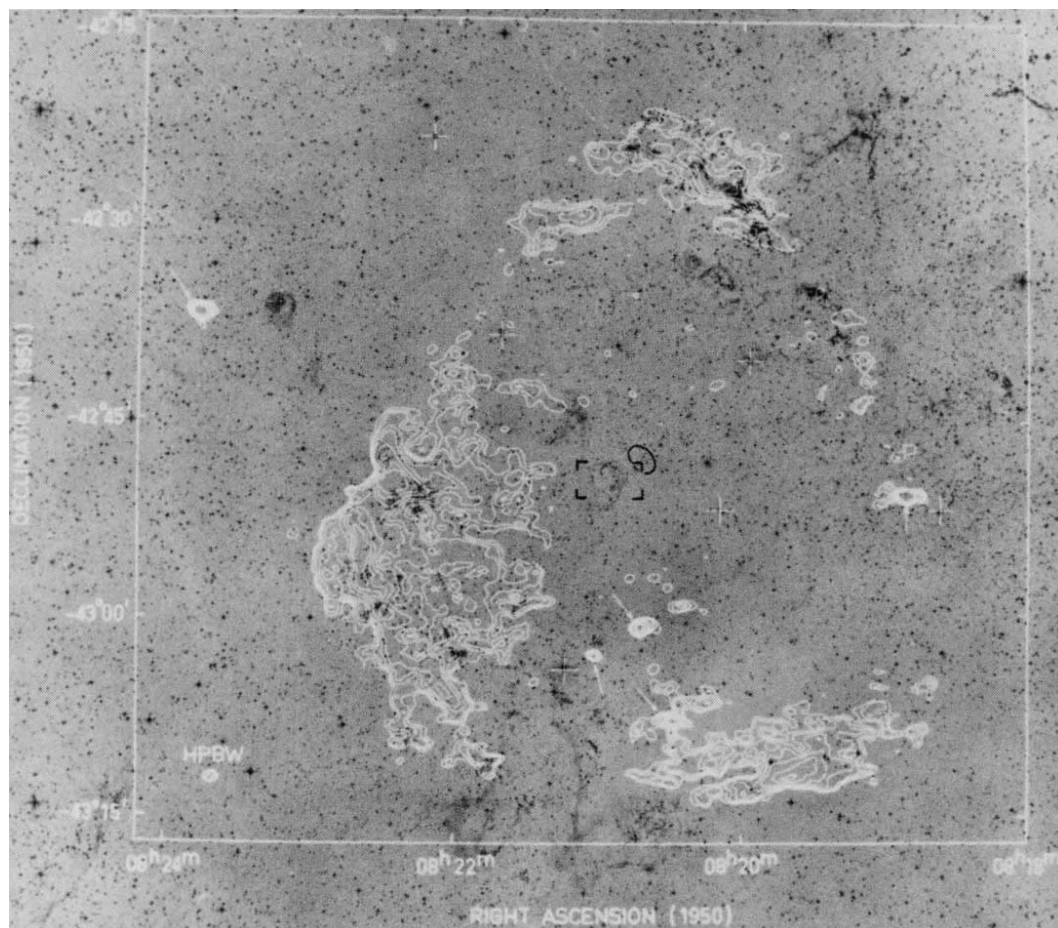
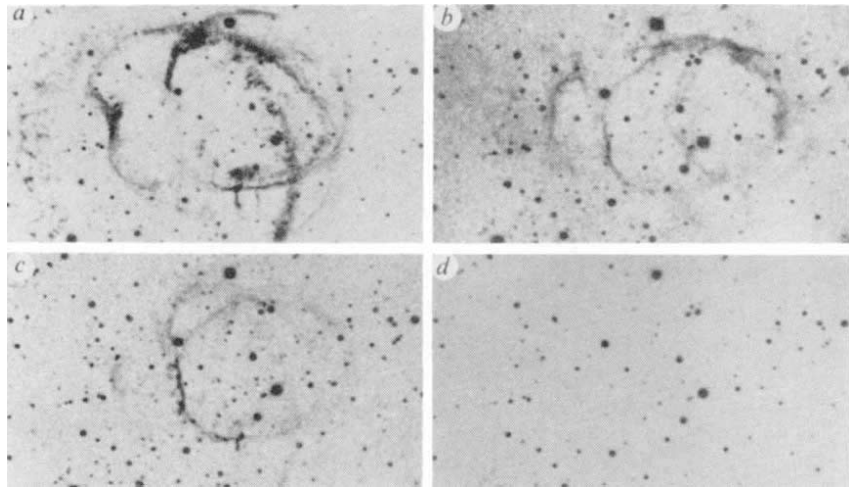


Fig. 1 Radio contours for Puppis A (ref. 2) superimposed on a deep $H\alpha + [\text{N II}]$ plate from the UK Schmidt Telescope¹. The swirl feature is visible as a group of faint filaments just east of the apparent centre. The 3×5 arcmin field for our CCD frames (Fig. 2) is indicated by the corners, and the expansion centre for the oxygen-rich filaments¹⁰ is shown by the black ellipse 3 arcmin to the north-west.

Fig. 2 Narrow-band CCD images of the Puppis A swirl region taken from the CTIO 4-m telescope in the light of $H\alpha + [N II]$ (a), $[O III]$ (b), $[S II]$ (c), and a 6,100-Å continuum band (d). The field is 3×5 arcmin; north is up and east is to the left.



single long-slit spectrum with an exposure time of 2,400 s at a position chosen to cross the brightest filaments on the $[O III]$ and $H\alpha + [N II]$ images. The spectra (Fig. 3) from these filaments are clearly different from those of shock-heated interstellar material and indicate high concentrations of heavy elements. The oxygen filament (Fig. 3a) shows strong lines of $[O I]$ and $[O II]$ in addition to $[O III]$, all at a radial velocity of -760 km s^{-1} . Lines of $[N II]$, $[Ne III]$ and faint Balmer lines all have the same radial velocity within the uncertainty of 50 km s^{-1} , which corresponds to about one quarter of the pixel width at 6,000 Å. This spectrum is similar to those of the 'omega' filament, of other fast, oxygen-rich knots in Puppis A (ref. 9), and of oxygen-rich filaments in young SNRs such as Cas A (ref. 11).

Table 1 CCD frames of Puppis A swirl

Ion	λ_c (Å)	$\Delta\lambda$ (Å)	Exposure (s)
$[O III]$	5,010	50	1,200
$H\alpha + [N II]$	6,560	60	600
$[S II]$	6,720	50	600
Continuum	6,100	150	300

Images taken using the CTIO 4-m telescope, 17 February 1988.

The spectrum of a bright filament from the $H\alpha + [N II]$ image (Fig. 3b) shows that most of the emission is in the $[N II]$ lines: the flux in $[N II]$ 6583 is seven times that in $H\alpha$. Although such high $[N II]/H\alpha$ ratios are quite rare in SNRs, they have been observed previously in bright filaments of Puppis A⁴⁻⁶. Nitrogen-rich material could have resulted from pre-supernova mass loss, similar to material in the quasi-stationary flocculi of Cas A (ref. 11) or as observed surrounding SN1987A (ref. 12). What is most unusual in the spectrum of Fig. 3b is a high velocity (-990 km s^{-1}) in a nitrogen-dominated filament. This may be a less extreme example of filaments such as the recently discovered 'fast-moving flocculi' in Cas A (ref. 13) which have been interpreted as material from the outer envelope of a Wolf-Rayet progenitor, ejected in the supernova event.

Another set of filaments in the swirl region has a flux in $[S II]$ up to four times higher than in $H\alpha + [N II]$, and ten times higher than in $[O III]$, as measured on the PFCCD frames of Fig. 2. The strongest $[S II]$ emission lies along an arc which continues to the north as a bright $H\alpha + [N II]$ filament. No spectra have yet been obtained to confirm that these features are $[S II]$ filaments, but the alternative interpretation, that they are redshifted $H\alpha + [N II]$, would require a radial velocity in excess of $5,000 \text{ km s}^{-1}$, over three times higher than any yet measured in Puppis A.

The swirl region comprises the most extraordinary confluence

of filaments anywhere in Puppis A, and among known SNRs it is matched only by Cas A in its variety of chemically peculiar debris. If these filaments are part of the Puppis A remnant, some highly unusual mechanism must be proposed to account for their unique morphology, chemistry and velocities. For example, we might be seeing an aneurysm of ejecta, which as it moves outward is confined to a low-density tunnel pointed in our direction, rather like a rifle barrel.

We wish, however, to suggest a more radical interpretation of the data, namely that the swirl region is the remnant of a second supernova which has taken place inside the Puppis A cavity more recently than the event that resulted in Puppis A itself. The evidence for a second supernova consists of the appearance of the optical filaments, the lack of corresponding radio or X-ray features, the unusual chemical composition and the timescale for the swirl expansion.

The morphology of the swirl region is that of a filamentary shell; in particular, the combination of $[O III]$ and $[S II]$ together form an almost complete elliptical shell $\sim 100'' \times 125''$ in size. The absence of associated radio emission rules out the possibility that this shell is simply a chance alignment of a background SNR. The surface brightness of the swirl region is certainly no higher than that of Puppis A itself², so its intrinsic diameter should, according to the Σ - D relation, be no smaller. This would require a distance $> 50 \text{ kpc}$, which is clearly unreasonable despite the uncertainty that attends Σ - D distances.

If the small shell is indeed an SNR, then its surroundings must be extremely unusual to obscure its presence both in the radio and in X-rays. The hot cavity inside an older SNR shell, like that of Puppis A, would provide a high-pressure, low-density environment which could render a second supernova remnant within the cavity undetectable at radio and X-ray frequencies. Instead of leading to a strong shock with associated radio and X-ray emission, most of the kinetic energy from the second supernova would probably re-energize the existing bubble¹⁴. A related model of a supernova within a cavity has been proposed to explain the X-ray morphology of the young SNR N132D in the Large Magellanic Cloud (LMC)¹⁵. Optical emission could result if dense knots of ejecta from the second supernova are heated conductively by the surrounding hot plasma. Conduction heating is especially important in metal-rich material, and the spectral characteristics would mimic those of shock excitation¹⁶.

The chemical composition and kinematics of the swirl structure are both indicative of a very young remnant. Its varied, metal-rich filaments are similar to those of Cas A, and the velocities are similar to those measured in other young, oxygen-rich SNRs such as N132D (ref. 17) and 0540-69 (ref. 18) in the LMC, and G292.0+1.8 (ref. 19) in our Galaxy. At the distance of Puppis A (2 kpc), the swirl radius of 1 arcmin is equivalent to 0.6 pc, and we have observed radial velocities $> 750 \text{ km s}^{-1}$

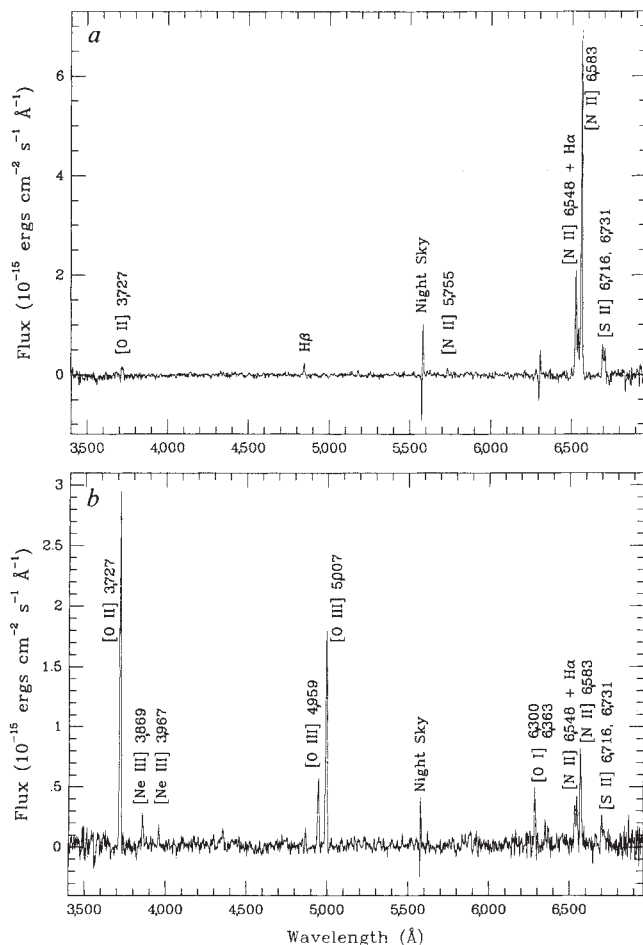


Fig. 3 Spectra of two knots from the same long-slit spectrum taken with the 2-D Frutti on the CTIO 4-m telescope. *a*, The brightest filament in the $H\alpha$ + $[N II]$ image is radiating primarily in $[N II]$ at -990 km s^{-1} ; $[S II]$ is present at the same velocity, and all the oxygen lines are extremely weak. *b*, The brightest knot in the $[O III]$ image shows lines of $[O I]$, $[O II]$ and $[O III]$, all at a radial velocity of -760 km s^{-1} .

for material near the rim. Assuming a comparable transverse velocity, we obtain a timescale of $<800 \text{ yr}$ for the swirl structure to expand to its present size. By contrast, the kinematic age for Puppis A based on proper motions of the oxygen-rich filaments⁹ is 3,700 yr, a value which is in reasonable agreement with age estimates based on the X-ray data.

Kinematic measurements provide potential tests for the 'second supernova' model. At 2 kpc a transverse velocity of 750 km s^{-1} would result in a proper motion of $0.08 \text{ arcsec yr}^{-1}$, so the timescale for the swirl expansion might be measurable within a few years. Meanwhile, more complete radial velocity data may reveal whether the swirl structure is made up entirely of (blue-shifted) material on the near side of the remnant, or whether it is indeed expanding.

Although the 'second supernova' model is speculative, it is interesting enough to merit conjecture about scenarios which could lead to two supernova events in rapid succession. One exotic possibility is that Puppis A was formed from a binary system in which the two components evolved in tandem to explode as near-simultaneous supernovae. The presence of filaments composed largely of oxygen and of oxygen-burning products indicates that Puppis A is the remnant of one or more highly evolved, massive stars, and the nitrogen-rich material suggests a Wolf-Rayet star in particular. Two nearly equal-mass stars in a massive binary system might evolve in tandem with

the exchange and loss of mass until the originally more massive one becomes a Wolf-Rayet star and then explodes as a supernova²⁰⁻²³. Mass gain by the original secondary would accelerate its evolution until it too becomes a supernova. If the masses are close enough, both components should explode as supernovae within a time interval that is short compared with the evolutionary timescale, possibly producing a remnant such as we observe in Puppis A. Although common models of massive binary systems^{24,25} typically assume an initial mass ratio of $\sim 2.5:1$ (so that the evolution of the two components proceeds at rates sufficiently different to produce long-lived X-ray binaries), components of nearly equal masses represent the most common configuration observed in unevolved massive binaries²⁶. Systems that are balanced delicately enough for both components to explode within a few thousand years would be rare, but possible. Regardless of how they might have occurred, two supernovae within a single shell could comprise the beginning of a super-shell, such has been proposed²⁷⁻²⁹ to explain the complex structure of the interstellar medium.

The swirl region in Puppis A is comprised of filaments that are spectroscopically, and probably chemically, more diverse than those in any other SNR except Cas A. All the optical properties—irregular shell morphology, peculiar chemistry and relatively high velocities—are characteristic of a young SNR, but the absence of significant radio or X-ray emission is most uncharacteristic. Although the swirl region may be an extraordinary feature of the well-known remnant, we prefer the suggestion that it signifies a second supernova inside Puppis A. Further spectroscopic observations, to investigate the chemical compositions and radial velocities of several filaments, are needed to confirm this interpretation. Proper motion measurements leading to a kinematic age for the swirl structure far younger than that of Puppis A itself might provide a definitive test.

We thank the support staff at CTIO for their assistance with the observations, and Chris Smith for help in preparing the figures. P.F.W., R.P.K. and J.P.H. are visiting astronomers at CTIO, which is operated by AURA, Inc. under contract from the US NSF. This work is supported by the NSF and NASA.

Received 22 August; accepted 11 November 1988.

1. Elliott, K. H., Goudis, C. & Meaburn, J. *Mon. Not. R. astr. Soc.* **175**, 605-611 (1976); Goudis, C. & Meaburn, J. *Astr. Astrophys.* **62**, 283-285 (1978).
2. Milne, D. K., Goss, W. M. & Danziger, I. J. *Mon. Not. R. astr. Soc.* **204**, 237-239 (1983).
3. Petre, R., Canizares, C. R., Kriss, G. A. & Winkler, P. F. *Astrophys. J.* **258**, 22-30 (1982).
4. Baade, W. & Minkowski, R. *Astrophys. J.* **119**, 206-214 (1954).
5. Danziger, I. J. & Dennefeld, M. *Publ. astr. Soc. Pacif.* **88**, 44-49 (1976).
6. Dopita, M. A., Mathewson, D. S. & Ford, V. L. *Astrophys. J.* **214**, 179-188 (1977).
7. Elliott, K. H. *Mem. Soc. Astr. Ital.* **49**, 477-479 (1978).
8. Shull, J. M. *Astrophys. J.* **269**, 218-228 (1983).
9. Winkler, P. F. & Kirshner, R. P. *Astrophys. J.* **299**, 981-986 (1985).
10. Winkler, P. F., Tuttle, J. H., Kirshner, R. P. & Irwin, M. J. in *Supernova Remnants and the Interstellar Medium, Proc. Int. astr. Un. Colloq. 101* (eds Roger, R. S. & Landecker, T. L.) 65-68 (Cambridge Univ. Press, 1988).
11. Chevalier, R. A. & Kirshner, R. P. *Astrophys. J.* **219**, 931-941 (1978).
12. Fransson, C. *et al. Astrophys. J.* (in the press).
13. Fesen, R. A., Becker, R. H. & Goodrich, R. W. *Astrophys. J.* **329**, L89-L92 (1988).
14. MacLow, M.-M. & McCray, R. *Astrophys. J.* **324**, 776-785 (1988).
15. Hughes, J. P. *Astrophys. J.* **314**, 103-110 (1987).
16. Shull, J. M. in *Supernova Remnants and the Interstellar Medium, Proc. Int. astr. Un. Colloq. 101* (eds Roger, R. S. & Landecker, T. L.) 161-178 (Cambridge Univ. Press, 1988).
17. Lasker, B. M. *Astrophys. J.* **237**, 765-768 (1987).
18. Mathewson, D. S., Dopita, M. A., Tuohy, I. R. & Ford, V. L. *Astrophys. J.* **242**, L73-L76 (1980).
19. Braun, R., Goss, W. M., Danziger, I. J. & Boksborg, A. in *Supernova Remnants and their X-ray Emission, Proc. Int. astr. Un. Symp. 101* (eds Danziger, J. & Gorenstein, P.) 159-164 (1983).
20. Willis, A. J. in *Interacting Binaries* (eds Eggleton, P. P. & Pringle, J. E.) 103-126, (Reidel, Dordrecht, 1984).
21. Abbott, D. C. & Conti, P. S. A. *Rev. Astr. Astrophys.* **25**, 113-150 (1987).
22. Massey, P. in *Wolf-Rayet Stars: Observations, Physics, Evolution* (eds de Loore, C. W. H. & Willis, A. J.) 251-262 (Reidel, Dordrecht, 1982).
23. de Loore, C. in *Wolf-Rayet Stars: Observations, Physics, Evolution* (eds de Loore, C. W. H. & Willis, A. J.) 343-358 (Reidel, Dordrecht, 1982).
24. van den Heuvel, E. P. J. in *Structure and Evolution of Close Binary Systems* (eds Eggleton, P., Mitton, S. & Whelan, J.) 35-61 (Reidel, Dordrecht, 1976).
25. Hellings, P. & de Loore, C. *Astr. Astrophys.* **161**, 75-84 (1986).
26. Garmany, C. D., Conti, P. S. & Massey, P. *Astrophys. J.* **242**, 1063-1076 (1980).
27. Cox, D. P. & Smith, B. W. *Astrophys. J.* **189**, L105-L108 (1974).
28. McKee, C. F. & Ostriker, J. P. *Astrophys. J.* **218**, 148-169 (1977).
29. McCray, R. in *Supernova Remnants and the Interstellar Medium, Proc. Int. astr. Un. Colloq. 101* (eds Roger, R. S. & Landecker, T. L.) 447-459, (Cambridge Univ. Press, 1988).

A sub-millimetre aperture synthesis array for nonsolar planet detection

Dayton L. Jones & David J. Diner

Jet Propulsion Laboratory, California Institute of Technology,
Pasadena, California 91109, USA

The detection of Earth-like planets orbiting other stars has great scientific interest, but is a challenging observational goal. At wavelengths shorter than $4\text{ }\mu\text{m}$ a terrestrial planet would be about 10^{10} times fainter than a solar-type star, and at a distance of 10 parsecs the planet-star separation would never exceed 0.1 arcsec. At wavelengths longer than $50\text{--}100\text{ }\mu\text{m}$ the brightness ratio is reduced by four to five orders of magnitude. Here we consider the possibility of using aperture synthesis to image other planetary systems at sub-millimetre wavelengths ($50\text{--}1,000\text{ }\mu\text{m}$). This approach requires only a small increase in dynamic range over that already demonstrated by the Very Large Array at longer wavelengths.

The concepts described here are part of a continuing study to determine which observational techniques could directly detect terrestrial and larger planets orbiting a star within 10 pc of the Sun, using no more than 24 h of observing time per star. This is a considerable challenge at any observing wavelength, but the nature of the challenge differs from radio to infrared and visible wavelengths. At radio wavelengths the primary challenge is obtaining sufficient sensitivity, whereas at infrared and visible wavelengths the primary challenge is obtaining sufficient dynamic range (that is, by suppressing instrumental sidelobes and scattered light for a direct imaging system, or nulling the central star for an infrared or visible interferometer¹⁻³).

Using wavelengths longer than $\sim 50\text{ }\mu\text{m}$ has the advantage that the ratio of thermal flux density from a solar-type star and a terrestrial-type planet is orders of magnitude smaller than at shorter wavelengths. In fact, the expected flux density ratio ($\sim 3 \times 10^5$) is only a few times the dynamic range of radio maps currently produced by the Very Large Array (VLA), suggesting that straightforward aperture synthesis could produce images of other planetary systems. At long (radio) wavelengths, however, the absolute flux density of both star and planet decreases rapidly with increasing wavelength, so the region near the peak of the planet's black-body emission should be optimal. Figure 1 shows the spectrum of a terrestrial planet and solar-type star. This letter considers the design of an aperture synthesis imaging instrument operating in this wavelength range (between 50 and $1,000\text{ }\mu\text{m}$).

The required instrument characteristics are an angular resolution of <0.05 arcsec, a total field of view $>2 \times 2$ arcsec, sufficiently high sensitivity to detect a terrestrial planet at 10 pc in 1 day, and a dynamic range sufficient to prevent sidelobes of the star from hiding a planet ~ 0.1 arcsec away. The ability to image the entire field of interest with full sensitivity at every pixel, and thus to map out the positions of all planets at once, is clearly important as we do not know how complex other planetary systems may be. Resolving planetary disks at such distances is not practical; for example, the angular size of the Earth at 10 pc is only 1 microarcsec.

The resolution and field of view requirements listed above imply baselines from 0.1λ to at least 4λ metres in length, where λ is the observing wavelength in μm . Thus a range of baseline spacings of at least 40 to 1 must be sampled by the array. The field of view also sets an upper limit of 0.1λ metres for the diameter of each individual telescope (in order that the primary beams will not be too small). Otherwise many observations of the region around a target star would be required, greatly reducing the efficiency of the search.

An additional limit imposed by the field of view is the maximum fractional bandwidth. For a heterodyne inter-

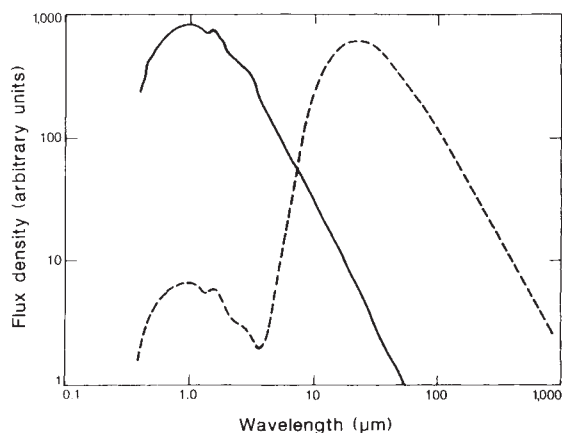


Fig. 1 The flux density of a terrestrial planet (dashed line) and a solar-type star (solid line) plotted against wavelength. The curve for the star has been multiplied by 10^{-8} . At wavelengths shorter than $\sim 4\text{ }\mu\text{m}$ the planet flux density comes from reflected starlight; at longer wavelengths it corresponds to thermal black-body emission.

ferometer it is useful to consider bandwidth in terms of frequency rather than wavelength. The relative response, R , to a source at an angular distance r from the centre of the field is

$$R = \frac{\sin(\pi \Delta\nu r / \nu\theta)}{(\pi \Delta\nu r / \nu\theta)}$$

where $\Delta\nu$ is the bandwidth, ν is the centre frequency of the observed radiofrequency band, and θ is the width of the synthesized beam. All sources not at the field centre will be radially elongated by $r \Delta\nu / \nu$. Preventing significant loss of sensitivity at the field edge requires that $\Delta\nu / \nu \ll \theta / r$. With $r = 2$ arcsec and a synthesized beam smaller than 0.05 arcsec, we need $\Delta\nu / \nu \ll 0.025$. For a multichannel system this limit applies to each individual intermediate frequency channel, not to the combined bandwidth.

The r.m.s. noise (S_n) in an image produced by aperture synthesis with all data weighted equally is given by

$$S_n = \frac{2^{1/2} k T_s}{\eta A [\tau \Delta\nu (N(N-1)/2)]^{1/2}} \text{ W m}^{-2} \text{ Hz}^{-1}$$

where the Boltzmann constant $k = 1.38 \times 10^{-23} \text{ W Hz}^{-1} \text{ K}^{-1}$, T_s is the system temperature, η is the aperture efficiency of the telescopes, A is the physical aperture of a single telescope, τ is the total integration time in seconds and N is the number of individual telescopes in the array.

The minimum total aperture of the array has been calculated assuming telescope efficiencies of 50%, system temperatures of $h\nu/k$ or 50 K, whichever was higher, a total IF bandwidth of 300 GHz (10% of the observing radiofrequency at $\lambda = 100\text{ }\mu\text{m}$), a total integration time of 1 day, and the use of global fringe fitting⁴ (see below). The large IF bandwidth could be produced by combining a large number of parallel channels, but the low system temperature will require significant advances in sub-millimetre mixer or amplifier technology.

The flux density S_p of a planet at temperature T subtending a solid angle of Ω steradians is

$$S_p = \frac{2h\nu^3\Omega}{c^2(e^{h\nu/kT} - 1)} \text{ W m}^{-2} \text{ Hz}^{-1}$$

For a terrestrial planet at 10 pc, $S_p \approx 7.3 \times 10^{-34} \text{ W m}^{-2} \text{ Hz}^{-1} = 7.3 \times 10^{-8} \text{ Jy}$ at $100\text{ }\mu\text{m}$. We would normally require $S_p/S_n \geq 5$, but if global fringe fitting is used the minimum useful signal-to-noise ratio can be decreased by an additional factor⁵ of $\sim \sqrt{N}$.

With these assumptions, the sensitivity requirement implies a total physical aperture NA (collecting area) of $\sim \lambda^2 \text{ m}^2$, where

λ in μm . This applies for wavelengths longer than $\sim 300 \mu\text{m}$. Because the collecting area of an individual telescope is proportional to λ^2 for a constant beamwidth, the total number of telescopes in the array remains constant ($N \approx 100$ in this example) for wavelengths longer than $\sim 300 \mu\text{m}$.

At shorter wavelengths the system temperature increases because of quantum effects ($T_{\text{sys}} \approx h\nu/k$), so the required collecting area decreases more slowly with decreasing wavelength. For wavelengths between ~ 50 – $300 \mu\text{m}$ the total aperture of the array is $\sim 300 \lambda^2 \text{ m}^2$. At wavelengths shorter than $\sim 20 \mu\text{m}$ the flux density of the planet turns over and the required aperture increases rapidly.

There is a minimum for the total aperture of $\sim 10^4 \text{ m}^2$ at $\lambda = 20$ – $30 \mu\text{m}$. Between ~ 9 and $100 \mu\text{m}$ the required aperture is within a factor of three of the minimum. This minimum is about 0.7 times the collecting area of the VLA. The detection of planets similar to Jupiter is much easier. Even with only one tenth as much total bandwidth (30 GHz), Jupiter could be detected at 10 pc in 24 hours with a total aperture of $\sim 2 \times 10^3 \text{ m}^2$ in the 20–50 μm region. This is only 15% of the VLA collecting area. Figure 2a shows the minimum collecting area as a function of wavelength for both terrestrial and Jupiter-type planets.

Figure 2b shows the number of individual telescopes needed for the array as a function of wavelength. At each wavelength, the telescopes are assumed to be as large as the field-of-view requirement allows. This figure makes it clear that the 20–50 μm spectral region is not desirable, despite the fact that it requires the smallest total collecting area. It may be more practical to use a wavelength closer to $300 \mu\text{m}$ to minimize the total number of telescopes rather than the total collecting area.

All of our sensitivity calculations assume that the total number of telescopes (N) in the array is large (given by the number of 0.1λ -metre diameter telescopes needed to reach the total required collecting area). For example, at $\lambda = 100 \mu\text{m}$ the maximum individual telescope diameter is 10 m and the total number of telescopes required to detect a terrestrial planet in 1 day is ~ 250 . For a Jupiter-like planet the number of telescopes required is ~ 40 , assuming a 1% bandwidth instead of 10%. There are several advantages to having a large N , including more accurate self-calibration of the individual telescope amplitudes and phases and, as mentioned before, an increase in the total sensitivity of the array by a factor of $\sim \sqrt{N}$ through global fringe fitting. However, if a smaller N is desired, the number of IF channels (and thus the total bandwidth) or the total observing time can be increased. As sensitivity is proportional to $\sqrt{\tau \Delta\nu}$ a large increase in τ or $\Delta\nu$ is required to significantly reduce the total collecting area. The maximum integration time is limited by the orbital motion of the planet around its parent star. In the case of a face-on planetary system, a terrestrial planet orbiting a star 10 pc away will move by half the synthesized-beam width in 2 weeks. Thus, if we are willing to accept some smearing of the planet image, the required number of telescopes (or total collecting area) could be reduced from the values given above by about a factor of four, at the expense of observing time.

The individual telescopes must not be so small that fringes cannot be detected on at least one baseline to each telescope. This suggests a hybrid array composed of many telescopes whose size is no larger than that set by the field-of-view requirement, along with one larger telescope whose primary beam may be too small to include any planets. The purpose of this single large telescope is to provide a more sensitive baseline to each of the other telescopes, so that global fringe fitting may be used. As described in detail by Schwab and Cotton⁴, global fringe fitting will then be able to detect the weaker fringes on all baselines between the small telescopes. This technique has come into use recently for very-long-baseline interferometry (VLBI) fringe searching.

The relatively large flux density of the star is an advantage here, as it allows the array to be phased up (focused) nearly continuously. By adjusting the individual telescope phases to

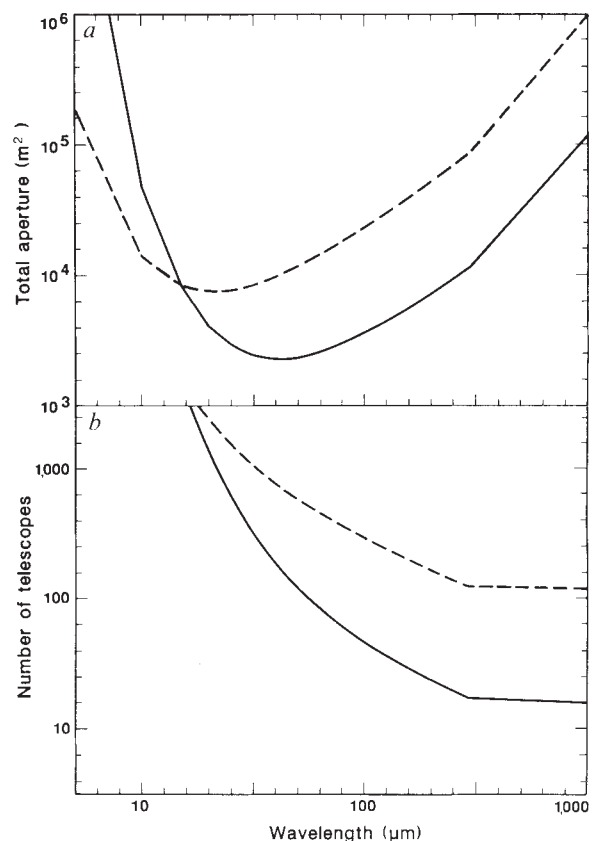


Fig. 2 Minimum total collecting area (*a*) and number of telescopes (*b*) for an aperture synthesis array designed to detect terrestrial planets at a distance of 10 pc in 24 hours. *a*, The dashed curve assumes a 300-GHz bandwidth, 50% aperture efficiency and a system temperature which is the higher of $h\nu/k$ or 50 K. Using wavelengths near 20–30 μm minimizes the required total physical size of the array. The solid curve shows the total aperture required to detect a gas giant planet similar to Jupiter, with the same assumptions as before except for a 30-GHz bandwidth. *b*, Number of telescopes required to provide the array collecting area. As above, the dashed curve is for a terrestrial planet and the solid curve is for a Jupiter-like planet. For wavelengths longer than $\sim 300 \mu\text{m}$ the number of telescopes is nearly constant, whereas for shorter wavelengths the number increases rapidly. Note that for the Jupiter-like planet only about 16 telescopes are needed at long wavelengths, assuming a bandwidth of 30 GHz.

keep the measured visibility phase from the star equal to zero on every baseline, the much weaker fringes from any planets will be nearly 'stopped' as well. This allows long coherent integrations.

The dynamic-range requirement implies that the array must have many separate baselines, and thus a large N , to provide dense and uniform coverage of the aperture (u, v) spectral plane. In addition, the (u, v) coverage should be 'centrally condensed'—that is, there should be more short baselines than long ones. This tends to reduce sidelobe levels in a manner similar to apodization in optical systems. A further requirement for very-high-dynamic-range-aperture synthesis mapping is an absence of errors that are baseline-dependent rather than telescope-dependent. This will require careful attention to polarization response, bandpass response and correlator design.

If we assume that the zodiacal light has a uniform surface brightness within the field of view, then its only effect is to increase the interferometer system temperature and thus the level of thermal noise in the images. The same is true of thermal emission from the telescope mirrors if the temperature remains constant. Both noise sources, if constant, are negligible com-

pared with the heterodyne receiver quantum noise. They may cause problems if spatial structure in the zodiacal light or telescope temperature fluctuations cause systematic variations in apparent fringe amplitude.

The zodiacal light surface brightness is given by

$$B_z = \frac{2h\nu^3 \epsilon}{c^2(e^{h\nu/kT} - 1)} \text{ W m}^{-2} \text{ Hz}^{-1} \text{ ster}^{-1}$$

For an emissivity ϵ of 3×10^{-7} , a temperature of 240 K and a wavelength of 100 μm , the surface brightness within a 2 arcsec field of view is about 5% of the flux density of a solar-type star at 10 pc. Thus, fluctuations in zodiacal light at the level of 1 part in 10^4 over sub-arcsecond angular scales, if present, must be discriminated from true planetary signals. Space-based visible or infrared mapping of selected regions of the zodiacal background with sub-arcsecond resolution would constitute a valuable precursor mission.

This project requires that a large number of small telescopes be deployed, which means that the cost per telescope must be small. In addition, the cross-correlation processing load will be very large. The correlator must be able to handle an extremely high data rate ($>10^{14}$ bits per second) because of the large number of baselines $N(N-1)/2$ and large bandwidth (many GHz). For comparison, the digital correlator at the VLA handles 351 baselines and a total bandwidth of 200 MHz. There has been considerable recent progress in the design of large digital

correlators, particularly in the development of the FX approach, in which the data from each telescope are Fourier-transformed to the frequency domain before cross-multiplication with all other spectra, which has made very large correlators more practical. Optical cross-correlation is a possibility, although baseline-dependent errors may be more difficult to control with an optical correlator.

A large interferometer array of this sort would be able to meet all of the goals listed above. It would also be an immensely valuable instrument for other areas of astronomical research. Studies of compact extragalactic radio sources, microwave background fluctuations, radio luminosity functions and source counts to very low flux densities are a few obvious examples. The practicality of this approach to nonsolar planet detection is dependent on financial, not basic technical, considerations.

We thank Robert Preston for helpful comments on an earlier version of this paper. This work was done at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

Received 8 September; accepted 14 November 1988.

1. Ken Knight, C. E. *Icarus* **30**, 422-433 (1977).
2. Bracewell, R. N. *Nature* **274**, 780-781 (1978).
3. MacPhie, R. H. & Bracewell, R. N. *Proc. SPIE* **172**, 271-278 (1979).
4. Schwab, F. R. & Cotton, W. D. *Astrophys. J.* **88**, 688-694 (1983).
5. Downes, D. *MM Array Memo 22* (National Radio Astronomy Observatory, Charlottesville, 1984).

Probable detection of organic-dust-borne aromatic C_3H_3^+ ions in the coma of comet Halley

A. Korth*, M. L. Marconi*, D. A. Mendis†, F. R. Krueger‡, A. K. Richter*, R. P. Lin§, D. L. Mitchell§, K. A. Anderson§, C. W. Carlson§, H. Rème||, J. A. Sauvaud|| & C. d'Uston||

* Max-Planck-Institut für Aeronomie, D-3411 Katlenburg-Lindau 3, FRG

† Department of Electrical Engineering and Computer Sciences, University of California, San Diego, California 92093, USA

‡ Ing.-Büro Krueger, D-6100 Darmstadt 12, FRG

§ Centre d'Etude Spatiale des Rayonnements, CNRS-Paul Sabatier University, F-31029 Toulouse, France

|| Space Sciences Laboratory, University of California, Berkeley, California 94720, USA

The heavy-ion analyser PICCA^{1,2} on the Giotto spacecraft was used to determine the composition and energy distribution of positively charged ions in the coma of comet Halley. Here we argue that the distinct peak observed at mass 39 AMU is due to the aromatic cation C_3H_3^+ (cyclopropenyl). Laboratory mass spectroscopy of unsaturated hydrocarbon-chain molecules and aromatics yields large intensities of mass-to-charge ratio (m/z) 39; the spectra of the aromatic molecules also indicate the presence of C_4H_3^+ ($m/z = 51$) and C_5H_5^+ ($m/z = 65$). These latter products are not seen in the PICCA data, suggesting that in the cometary environment unsaturated hydrocarbon chains predominate over aromatics. We suggest that the progenitor of the hydrocarbon chains (which are apparently the parents of the C_3H_3^+ ion) are the 'CHON' dust particles observed around the comet³; the C_3H_3^+ ion may also be one of the parent molecules of C_3 , which is a common feature of cometary spectra⁴, although its origin has remained obscure.

The PICCA (positive-ion-cluster-composition analyser) instrument uses an electrostatic analyser and measures the ratio

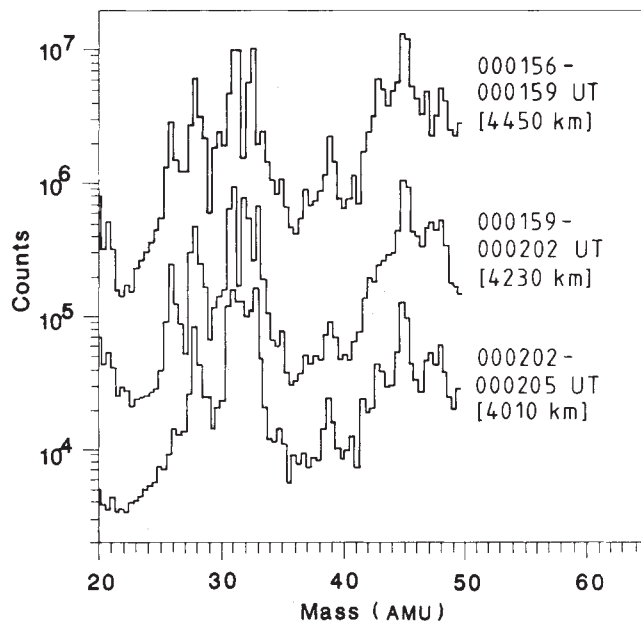


Fig. 1 Three consecutive measurements of mass spectra. The time is given in spacecraft event time and corresponds to cometocentric distances of (top to bottom) 4,450, 4,230 and 4,010 km.

of energy to charge (E/z). It takes advantage of the large relative fly-by velocity and the fact that cometary ions in the inner coma have very low temperatures (thermal velocity $\ll$ ram velocity) and are invariably singly charged. Thus the measurements of E/z are, to a very good approximation, proportional to the mass. The mass resolution is $\Delta m = 0.4$ AMU (atomic mass units) in the mass range 10-50 AMU, and $\Delta m = 1$ AMU in the mass range 51-210 AMU. Ions were detected by two channel electron multipliers with different geometric factors to provide a large dynamic range. The spectrum integration time of 3.2 s is different from the spacecraft spin period of 4 s.

In the following, we show spectra from the PICCA instrument

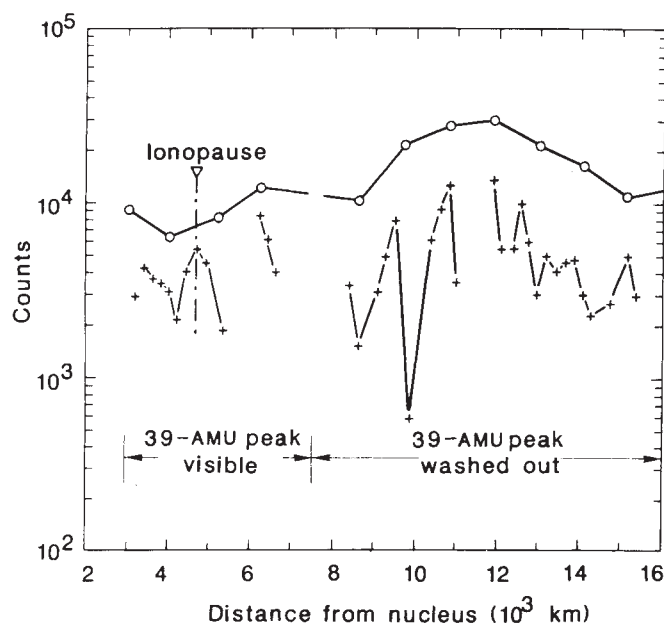


Fig. 2 Radial dependence of the mass channels 39 AMU (+) and 63–73 AMU (O). The intensity, given in counts s^{-1} , is proportional to the density. The position of the ionopause is indicated.

inside the cometary ionopause. The ionopause (which is a tangential discontinuity surface) is the outer boundary of the magnetic-cavity region detected at a cometocentric distance of $\sim 4,600$ km during Giotto's inbound passage⁵. In this region, the ion temperatures are as low as 300 K (ref. 6). In this paper, we will discuss peaks in the spectra at masses of 37, 39 and 41 AMU. Although the PICCA instrument measured ions up to 105 AMU, only the part of the spectra close to the peaks of interest are presented here.

In Fig. 1, three consecutive spectra are shown, displaced vertically for clarity. The time is given in spacecraft event time for each spectrum and corresponds to a certain cometocentric distance. The distance from the comet varies from $\sim 4,450$ km to $\sim 4,010$ km. The count rates correspond to the central spectrum and are proportional to the density. The two other spectra are shifted up and down by half an order of magnitude relative to the central one. The peak at mass 39 AMU which is visible in all three of the displayed spectra can be observed continuously up to cometocentric distances of $\sim 7,500$ km with no dramatic change at the ionopause. Outside the ionopause, the thermal-energy spread of the ions limits the effective mass resolution of the instrument so that the peaks get washed out. The transition from sharp to washed-out is not accompanied by a pronounced drop in the count rate (density).

Figure 2 shows the radial dependence of the 39-AMU peak intensity, which is proportional to the density. The density of the 39-AMU species inside the ionopause at $\sim 4,200$ km is ~ 5 ions per cm^3 . The values of the intensity for a cometocentric distance $r > 7,500$ km are obtained by the fitting process described by Mitchell *et al.*⁷. For comparison, a solid line trace is given which shows the count rates summed from the channels 64–73 AMU (inclusive) and averaged over 16 s. Both curves show similar features, including an intensity maximum between 11,000 and 12,000 km. Unfortunately, it is not possible to deduce a reliable radial profile because of the spin modulation of the instrument and data gaps.

The unambiguous identification of a given peak in the spectra shown in Fig. 1 with a particular ion is not possible. In Table 1, however, we have listed the dominant ion species that are expected from a CHON-dominated medium⁸. As pointed out earlier, we take $z = 1$ (singly charged ions). The PICCA data

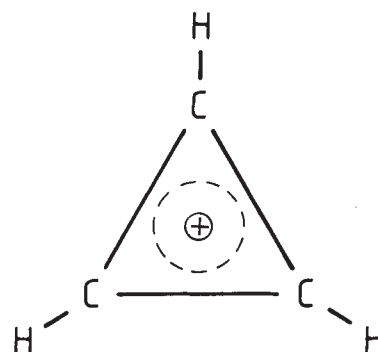


Fig. 3 The structure of $C_3H_3^+$, the cyclopropenyl cation.

show a strong preference for odd mass numbers over even mass numbers; this distribution is typical for a nitrogen-poor non-radical ion mixture. Radical ion species are thus unlikely and are shown in parentheses in Table 1.

The most likely identification of the $m/z = 39$ feature is as the aromatic $C_3H_3^+$ (cyclopropenyl) rather than as a hydrocarbon-chain molecule (see below). The cyclopropenyl ion has a ring structure and shows an aromatic character, having an appropriate number of π electrons (Fig. 3). It has no permanent electric dipole moment and will not be visible in the microwave spectrum.

Although aromatics have a highly unsaturated character, they do not enter into addition reactions and are therefore very stable. Besides $C_3H_3^+$ there are only a few other stable cations with aromatic character. For example $C_7H_7^+$, tropylium ($m/z = 91$), is possibly one of the ions that we identified in the PICCA instrument outside the ionopause.

Inside the ionopause individual mass peaks have been resolved below 70 AMU, whereas outside the ionopause a number of broad mass groups have been observed, one of which is centred at 90 AMU (ref. 9). Contributions to this broad mass group need not come only from an ion of $m/z = 90$, but may also be from ions of neighbouring masses ($m/z = 89$ (dehydrotropylium) and $m/z = 91$ (tropylium)). Although Mitchell *et al.*⁹ suggest that $m/z = 90$ corresponds to the ion $HOCH_2CH_2COOH^+$, observed in laboratory residues resulting from simulation experiments¹⁰, $m/z = 91$ could correspond to $C_7H_7^+$, which is known to be very abundant and stable. Recently, Giard *et al.*¹¹ made the first identification of polycyclic aromatic hydrocarbon molecules in the diffuse emission of the galactic disk.

Laboratory mass spectroscopy of the destruction products of unsaturated hydrocarbon-chain molecules and aromatics shows a pronounced peak at $m/z = 39$ (refs 12, 13). Aromatics also commonly show products at $m/z = 51$ and $C_4H_3^+$ 65. These products, which could be identified as $C_4H_3^+$ and $C_5H_5^+$, are not seen abundantly in the PICCA data. On the other hand, unsaturated hydrocarbon-chain molecules or compounds with oxygen do not show abundant peaks at the above values but instead show peaks at $m/z = 41$ (possibly the cyclopropanyl ion, $C_3H_5^+$) and at $m/z = 43$. The PICCA instrument sees these latter ions rather abundantly. We therefore conclude that the main progenitor of the $m/z = 39$ peak is a long unsaturated chain

Table 1 Possible identification of mass peaks in a CHON-dominated medium

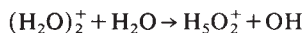
m/z	Ion species
37	C_3H^+ , $H^+(H_2O)_2$
39	$C_3H_3^+$, (C_2NH^+)
41	$C_3H_5^+$, C_2HO^+ , $(C_2H_3N^+)$

Radicals are enclosed in parentheses.

molecule, although aromatic progenitors are not entirely excluded.

As indicated in Table 1, the 39-AMU mass peak could also be produced by the molecule C_2NH^+ , assuming only CHON as building blocks. This molecule is a radical, however, and no indications of other homologous radical ions are found in the spectra. Moreover, its reactivity, relative to $C_3H_3^+$, is much higher and therefore there is a high probability of reactions with H_2O , which is the most dominant molecule in the inner coma. The possibility that $m/z = 39$ could also correspond to potassium is ruled out because there is far too little K^+ in the dust to produce such an abundant yield of gas-phase K^+ (ref. 14).

A further decomposition of mass 39 by dehydrogenation to C_3H^+ (mass 37, which is also found, to a lesser extent, in the PICCA spectra (see Fig. 1)) is only possible in high-energy collision processes. These are not likely in an inner cometary coma, where the plasma is very cold (≤ 300 K). But we note that a significant contribution to the mass-37 peak could also come from $H_3O_2^+$ by the reaction:



if a substantial fraction of water sublimating from the nucleus is in the form of the dimer $(H_2O)_2$ (ref. 15).

We now consider the origin of $C_3H_3^+$. Is it derived from the dust component or from some gas component emanating from the nucleus? Assuming that the ion of mass 39 participates readily in the atmospheric chemistry, the required abundance of the progenitor molecule can be estimated. The primary gas-phase destruction processes in the cometary atmosphere for any given ion are ion-molecule reactions with H_2O , proceeding at a typical rate $\tau_{im} \approx 10^{-9} n_{H_2O} s^{-1}$, and recombination, with a rate $\tau_{rec} \approx 10^{-6} (300/T_e)^{1/2} n_e s^{-1}$ (ref. 16). Using a number density for H_2O of $n_{H_2O} = 5 \times 10^6 cm^{-3}$ (ref. 17), electron temperature $T_e = 100$ K and electron number density $n_e = 1,500 cm^{-3}$ (from model calculations)¹⁸ at a distance of $3-4 \times 10^3$ km, $\tau_{im} = 5 \times 10^{-3} s^{-1}$ and $\tau_{rec} = 2.5 \times 10^{-3} s^{-1}$. This implies that the ion $m/z = 39$ must be produced at a rate of $\sim 7 \times 10^{-3} n_{39^+} cm^{-3} s^{-1}$ (where n_{39^+} is the number density of the ion of mass 39 and $\tau_{rec} + \tau_{im} \equiv \tau_d \approx 7 \times 10^{-3} s^{-1}$ is the total destruction rate).

The production of the $m/z = 39$ ion may occur through either (1) ion-molecule reactions ($X^+ + Y \rightarrow X_{39}^+ + \text{products}$); (2) photoionization of a neutral ($X + h\nu \rightarrow X_{39}^+ + \text{products}$); or (3) photodissociation of an ion ($X^+ + h\nu \rightarrow X_{39}^+ + \text{products}$). Clearly, (1) requires that Y be about as abundant as H_2O because X_{39}^+ is destroyed by H_2O by the analogous reaction. Reaction (2) requires n_x (the number density of X) to be equal to $n_{39^+} (\tau_d/\tau_{pi})$, where τ_{pi} is the photoionization rate. If $\tau_{pi} = 10^{-6} s^{-1}$ (a typical photoionization timescale) and $n_{39^+} = 5 cm^{-3}$, then $n_x = 3.5 \times 10^4 cm^{-3}$, which is a density comparable to that of the species, such as NH_3 and CH_4 , that are the next most abundant after H_2O and CO. Again, n_x appears to be too large to remain unobserved. Reaction (3) requires $n_{X^+} = n_{39^+} (\tau_d/\alpha)$, where α is the photodissociation rate of X^+ . To estimate α we consider the most steeply decreasing mass peak observed between 3,000 km and 4,000 km, which is the mass-75 group, and attribute the entire decrease to photo-processes. Assuming that the ion velocity inside the ionopause is equal to the neutral velocity (a good approximation because of the strong collisional coupling between ions and neutrals there) and taking the latter to be $\approx 0.8 km s^{-1}$ (ref. 19), we get $\alpha \leq 5 \times 10^{-4} s^{-1}$, and consequently, $n_{X^+} \geq 14 n_{39^+}$. But the only observed mass peak in the PICCA data beyond mass 39 with a significantly higher intensity is that for $m/z = 45$ (which is ~ 5 times higher). This is unlikely to produce the $m/z = 39$ ion through reaction (3).

Although the above gas-phase calculations are based on reaction rates that are somewhat uncertain, we believe that they are sufficiently reliable for our conclusions to be valid. Thus we can eliminate a gas-phase progenitor molecule for $C_3H_3^+$, so that the only alternative source that could be significant seems to be the observed circumnuclear CHON dust particles. This is not sur-

prising, as dust has already been strongly implicated as a source of CO (which is the second most abundant species after H_2O , comprising 10–20% of the cometary atmosphere), as well as a source of CN.

This work was supported by the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, by the Bundesministerium für Forschung und Technologie, by CNES and by NASA.

Received 19 September; accepted 10 November 1988.

1. Korth, A. *et al. Nature* **321**, 335–336 (1986).
2. Korth, A. *et al. J. Phys. E* **20**, 787–792 (1987).
3. Kissel, J. *et al. Nature* **321**, 280–282 (1986).
4. A'Heam, M. F. in *Comets* (ed. Wilkening, L. L.) 448 (Univ. of Arizona Press, Tucson, Arizona, 1982).
5. Neubauer, F. M. *Astr. Astrophys.* **187**, 73–79 (1987).
6. Schwenn, R. *et al. Astr. Astrophys.* **187**, 160–162 (1987).
7. Mitchell, D. L. *et al. Eur. Space Ag. SP-250*, 203–205 (1986).
8. Silverstein, R. M., Bassler, G. C. & Morrill, T. C. *Spectroscopic Identification of Organic Compounds* 41 (Wiley, New York, 1974).
9. Mitchell, D. L. *et al. Adv. Space Res.* (in the press).
10. Schutte, W. A. thesis, Univ. of Leiden (1988).
11. Giard, M. *et al. Astr. Astrophys.* **201**, L1–L4 (1988).
12. Spiteller, G. *Massenspektrometrische Strukturanalyse organischer Verbindungen* (Verlag Chemie, Weinheim, 1966).
13. McLafferty, F. W. *Interpretation of Mass Spectra*. (Univ. Science Books, Mill Valley, California, 1980).
14. Kissel, J. & Krueger, F. R. *Nature* **326**, 755–760 (1987).
15. Murad, E. & Bochsler, P. *Nature* **326**, 366–367 (1987).
16. Mendis, D. A., Houppis, H. L. F. & Marconi, M. L. *Fundam. cosmic Phys.* **10**, 1–380 (1985).
17. Krankowsky, D. *et al. Nature* **321**, 326–329 (1986).
18. Marconi, M. L. & Mendis, D. A. *Astrophys. J.* **330**, 513–517 (1988).
19. Lämmerzahl, P. *et al. Astr. Astrophys.* **187**, 169–173 (1987).

Estimation of the rate of volcanism on Venus from reaction rate measurements

Bruce Fegley Jr & Ronald G. Prinn

Department of Earth, Atmospheric and Planetary Sciences,
Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139, USA

Maintenance of the global H_2SO_4 clouds on Venus requires volcanism to replenish SO_2 , which is continually removed from the atmosphere by reaction with calcium minerals on the planet's surface^{1–4}. Here we present laboratory rate data for the reaction between SO_2 and calcite ($CaCO_3$) to form anhydrite ($CaSO_4$). If this reaction rate represents the SO_2 reaction rate on Venus, then all SO_2 in the venusian atmosphere (and thus the clouds) will disappear in 1.9 Myr unless volcanism replenishes the lost SO_2 . The required volcanism rate, which depends on the sulphur content of the erupted material, is in the range 0.4–11 km^3 of magma erupted per year. The Venus surface composition at the Venera 13, 14 and Vega 2 landing sites^{5,6} implies a volcanism rate of approximately 1 $km^3 yr^{-1}$. This geochemically estimated rate can be used to determine if either (or neither) of two discordant geophysically estimated rates (2 $km^3 yr^{-1}$ versus 200–300 $km^3 yr^{-1}$)^{7–9} is correct. It also suggests that Venus may be less volcanically active than the Earth.

Calcite acts as a net sink for SO_2 on Venus by incorporating it into the crust through the net reaction^{1–4}



which proceeds to the right because the observed SO_2 abundance on Venus is ~ 100 times higher than the value in equilibrium with coexisting calcite and anhydrite⁴. Calcite is also predicted by several authors to be present on the Venus surface because it appears to be capable of buffering the CO_2 abundance on Venus through the Urey reaction ($CaCO_3 + SiO_2 = CaSiO_3 + CO_2$)^{10–13}. These predictions concerning possible roles for calcite on Venus, the exemplary nature of reaction (1), and the availability of high-purity $CaCO_3$ in several forms led us to study reaction (1) as an example of SO_2 removal processes. However, as

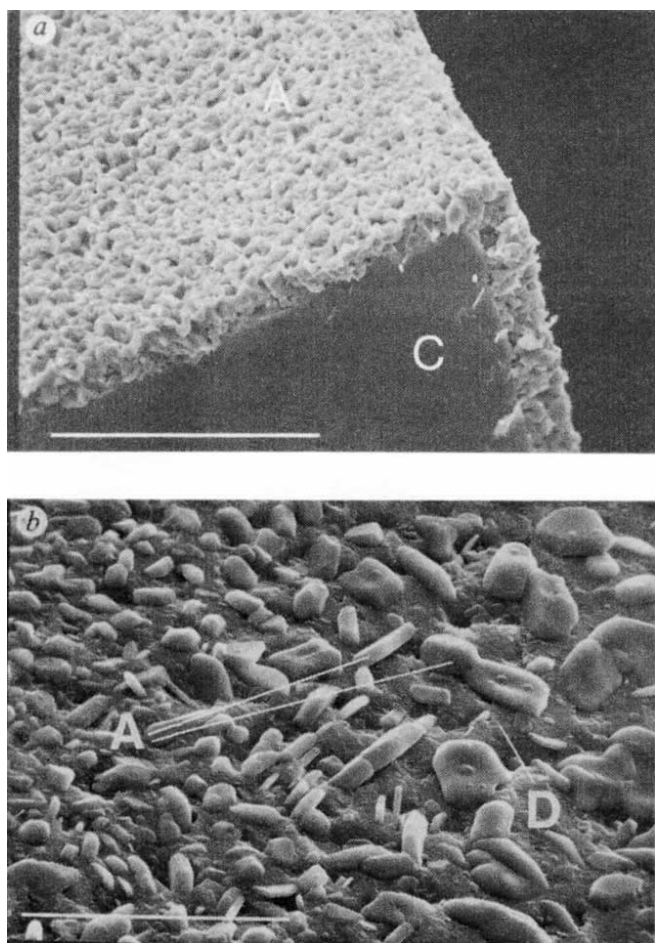


Fig. 1 Scanning electron micrographs of chemically reacted samples. *a*, Fracture surface, showing a CaSO_4 (anhydrite) layer (A) completely covering a partially reacted calcite crystal (C) (850 °C for 192 h). *b*, Top surface, showing CaSO_4 (A) growing on a partially reacted diopside crystal (D) (833 °C for 48 h). Scale bars indicate 100 μm (*a*) and 50 μm (*b*).

Fig. 1 illustrates, SO_2 also reacts with other calcium minerals (such as, diopside, $\text{CaMgSi}_2\text{O}_6$), which are believed to be present on the surface of Venus^{14,15}. The kinetics for these alternative reactions are also being measured and will be reported later.

Reaction (1) was studied by heating clear calcite crystals (Iceland spar) of known impurity content (~ 350 p.p.m. total impurities from analysis by inductively coupled plasma emission spectroscopy), weight and surface area in SO_2 -bearing gas streams for varying time periods at three different temperatures. Experiments were done at temperatures of 600–850 °C and ambient atmospheric pressures with SO_2 - CO_2 gas mixtures nominally containing 1% (by volume) SO_2 . Thus, the molecular number density of SO_2 is about the same as that at the surface of Venus, where the atmospheric pressure is ~ 100 times larger but the SO_2 volume mixing ratio is ~ 100 times smaller¹⁶. The mixtures also typically contained <1 p.p.m. O_2 and ≤ 10 p.p.m. H_2O as impurities. Experiments were also done with gas mixtures containing up to 100 p.p.m. O_2 ; however, the rates were independent of oxygen fugacity (and thus of the CO/CO_2 ratio) within this range of O_2 content. Standard techniques were used to monitor and control gas flow and temperature (to better than ± 5 °C). The rate of reaction (1) was determined by three independent methods: (1) measuring the weight gain at the end of reaction, (2) using scanning electron microscopy (SEM) to measure the porosity and thickness of the reacted surface layers from micrographs of external and fracture surfaces (see, for

example, Fig. 1), and (3) SO_4^{2-} analyses of reacted samples by ion chromatography. SEM examination of many samples (in combination with X-ray line scans and energy-dispersive spectroscopy (EDS) on the SEM) showed that CaSO_4 layers covered all external surfaces of the reacted samples. X-ray powder patterns of samples and standards confirmed that the layers were anhydrite (in particular, there was no evidence for calcium sulphite) and that unreacted calcite was the only other phase present in the samples. The ion chromatography also verified the absence of anion impurities at levels ≥ 10 p.p.m.

The experimental run conditions and rate data are summarized in Table 1 and Fig. 2. Rate data using the weight-gain, SEM and SO_4^{2-} analysis methods for the same sample agree within the combined experimental uncertainties (1σ) shown in Table 1. A weighted linear least-squares fit to the data gives a rate equation $R = 10^{19.64(\pm 0.28)} \exp(-15,248(\pm 2,970)/T)$ molecules $\text{cm}^{-2} \text{s}^{-1}$, with a corresponding activation energy of $126.8 \pm 5.4 \text{ kJ mol}^{-1}$. The rate of reaction (1) was extrapolated downward to Venus surface temperatures (~ 660 – 750 K) using this equation and the Pioneer Venus radar-altimetry data¹⁷ and atmospheric (P-T) profile¹⁸ to take into account the altitude dependence of the rate. The derived global-mean SO_2 reaction rate is $\sim 4.6 \times 10^{10}$ molecules $\text{cm}^{-2} \text{s}^{-1}$, equivalent to $\sim 1 \mu\text{m}$ of CaSO_4 being deposited per year. However, aeolian weathering¹⁹ by the $\sim 1.0\text{-m s}^{-1}$ winds on Venus²⁰ will probably remove this thin CaSO_4 layer on a shorter timescale, so the laboratory rate data for reaction (1) are applicable to Venus.

Chemical analyses of the Venus surface by the Venera 13, 14 and Vega 2 spacecraft^{5,6} give a CaO content (by mass) of $7.90 \pm 0.51\%$ (weighted mean $\pm 1\sigma$); this is also taken as the areal percentage because of similar rock and oxide densities. Assuming that the experimental rate is representative of the rate at which SO_2 is depleted by reaction with calcium minerals on the Venus surface, the observed SO_2 column density (2.2×10^{23} molecules cm^{-2})¹⁶ would be removed from the Venus atmosphere in $\sim 1.9 \times 10^6$ yr in the absence of a comparable sulphur source. However, maintenance of the global H_2SO_4 clouds, which are formed by the ultraviolet-sunlight-powered conver-

Table 1 Rate data used to calculate the activation energy

Sample	T (°C)*	Duration (hours)	Rate $\times 10^{-12}$ (molecules $\text{cm}^{-2} \text{s}^{-1}$)†
R44-1	602	95	$1.67 \pm 0.19\ddagger$
R45-1	602	314	0.85 ± 0.77
			$1.05 \pm 0.12\ddagger$
R45-3	602	314	$1.01 \pm 0.11\ddagger$
R45-4	602	314	1.20 ± 0.51
			$0.68 \pm 0.48§$
R36-1	746	96	8.43 ± 2.93
			$13.6 \pm 6.0§$
R36-2	746	96	23.2 ± 6.2
			$19.9 \pm 11.0§$
R50-1	747	192	9.19 ± 1.60
R50-3	747	192	10.5 ± 7.8
R51-1	748	96	13.2 ± 4.0
R51-2	748	96	10.3 ± 5.6
R43-1	848	91	60.4 ± 6.8
			$57.1 \pm 31.7§$
R43-2	848	91	53.2 ± 7.0
			$67.2 \pm 43.4§$
R43-3	848	91	58.1 ± 7.0
			$48.4 \pm 12.1§$

* The estimated temperature uncertainty is ± 5 °C.

† The number of SO_2 molecules reacted per cm^2 surface area per second. The rate is from weight gain measurements unless noted otherwise. The 1σ uncertainties are also listed.

‡ The rate from ion chromatography analyses of SO_4^{2-} content of reacted sample

§ The rate from scanning electron microscope (SEM) measurements of anhydrite layer thickness and porosity.

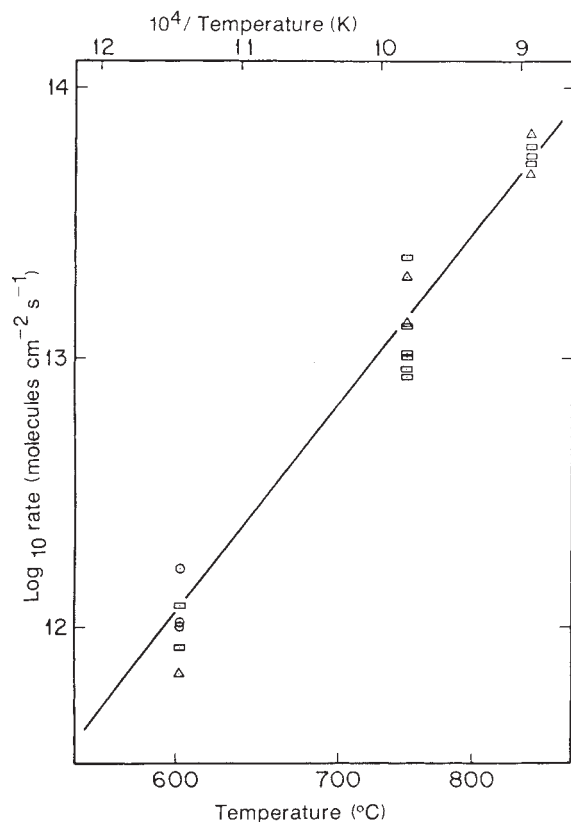


Fig. 2 Arrhenius plot of rate data from Table 1 against $1/T$. The line is a weighted linear least-squares fit, and indicates an activation energy of $126.8 \pm 5.4 \text{ kJ mol}^{-1}$. Rates determined from independent methods (weight gain (rectangles), SEM (triangles), SO_2 analyses (circles)) for the same sample agree within the 1σ experimental uncertainties.

sion of SO_2 into H_2SO_4 cloud particles²¹, requires a comparable sulphur source. This source must be endogenic, because infalling cosmic material cannot supply enough sulphur. The calculated SO_2 removal rate on Venus is equivalent to $\sim 2.8 \times 10^{13} \text{ g sulphur per year}$. However, the measured terrestrial flux of infalling material, which is roughly applicable to Venus because of similar sizes and masses, is only $\sim 7.8 \times 10^{10} \text{ g yr}^{-1}$ (ref. 22), or only $\sim 4.6 \times 10^9 \text{ g S yr}^{-1}$ assuming the most sulphur-rich CI chondrite composition²³.

The most plausible endogenic source is volcanism, which has occurred on Venus in the past²⁴, and which may have led to increased SO_2 levels above the Venus cloud-tops observed by the PV orbiter^{25,26}. The rate of volcanism required to balance SO_2 depletion by reactions with calcium minerals on the Venus surface depends on the sulphur content of the erupted material (gas and magma). If this material has an overall sulphur/silicon mass ratio of 0.03 (the weighted mean of the Venera 13, 14 and Vega 2 surface analyses^{5,6}), the corresponding rate of volcanism is $\sim 1 \text{ km}^3 \text{ yr}^{-1}$. Figure 3 illustrates that this rate is ~ 20 times smaller than the terrestrial value^{8,27}. Two other plausible models for the S/Si ratio of the erupted material on Venus are also shown: S/Si ≈ 0.1 , as in ordinary chondrites²³, giving a rate of $\sim 0.4 \text{ km}^3 \text{ yr}^{-1}$ and S/Si ≈ 0.004 , as in the Earth's crust²⁸, giving a rate of $\sim 11 \text{ km}^3 \text{ yr}^{-1}$. All rates will scale as ϵ^{-1} , where ϵ is the sulphur degassing efficiency for erupted material ($0 < \epsilon \leq 1$). However, ϵ is likely to be close to unity because sulphur gases are expelled directly into the atmosphere, and hot, reactive sulphur-bearing magmas will rapidly lose sulphur to the atmosphere. Our experiments on the chemical weathering by CO_2 of a possible magmatic sulphur-bearing mineral pyrite FeS_2 indicate complete loss of sulphur in hours at sub-magmatic temperatures ($\sim 980^\circ\text{C}$).

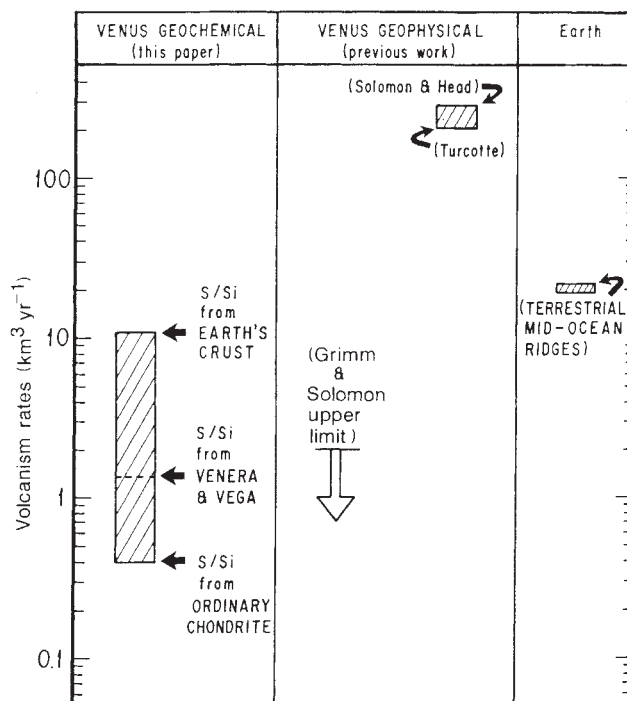


Fig. 3 Geochemical and geophysical estimates of the rate of volcanism on Venus, compared with the terrestrial rate. Three plausible models for the S/Si ratio of erupted material on Venus yield rates of ~ 0.4 to $\sim 11 \text{ km}^3 \text{ yr}^{-1}$ (see text). Adopting the Venus surface S/Si ratio from Venera 13, 14 and Vega 2 analyses^{5,6} yields a rate of $\sim 1 \text{ km}^3 \text{ yr}^{-1}$, or about 20 times less than the rate of volcanism on Earth. The terrestrial rate of $\sim 20 \text{ km}^3 \text{ yr}^{-1}$ is based on a plate creation rate²⁷ of $\sim 3 \text{ km}^2 \text{ yr}^{-1}$ and a crustal thickness of $\sim 6 \text{ km}$.

Figure 3 also illustrates that widely discordant volcanism rates have been estimated by two different geophysical methods. Rates of $\sim 200\text{--}300 \text{ km}^3 \text{ yr}^{-1}$ are estimated by scaling terrestrial heat production to Venus and transporting the outward heat flux solely by volcanism^{8,9}, and an upper limit of $2 \text{ km}^3 \text{ yr}^{-1}$ is estimated by analysis of impact-crater areal densities on Venera 15, 16 radar images⁷. The present, geochemically derived volcanism rate is independent of these two methods and can be used, in principal, to test if either (or neither) is correct. In particular, the values computed by the former geophysical method do not agree with our geochemical values. This work also suggests that on average Venus may be less volcanically active than the Earth. Finally, the geochemically estimated volcanism rate can also be used to constrain water loss on Venus through oxidation of FeO-bearing basalt³. This work is in progress and will be reported elsewhere.

We thank W. Zamechek for chemical analyses, and L. Rigione, S. Rowell and A. Petropoulos for assistance in the laboratory. This work was supported by the NASA Innovative Research Program. The MIT UROP program also partially supported S.R. and A.P.

Received 12 July; accepted 21 November 1988.

1. Prinn, R. G. in *Recent Advances in Planetary Meteorology* (ed. Hunt, G. E.) 1–15 (Cambridge University Press, London, 1985).
2. Prinn, R. G. *Sci. Am.* **252**, 46–53 (1985).
3. Prinn, R. G. in *The Photochemistry of Atmospheres* (ed. Levine, J.) 281–336 (Academic, New York, 1985).
4. Fegley, B. Jr *Lunar planet. Sci.* **XIX**, 315–316 (1988).
5. Surkov, Yu. A., Barsukov, V. L., Moskal'yova, L. P., Kharyukova, V. P. & Kemurdzhian, A. L. *Proc. 14th, Lunar planet. Sci. Conf. J. geophys. Res.* **89**, B393–B402 (1984).
6. Surkov, Yu. A. *et al. Proc. 17th, Lunar planet. Sci. Conf. J. geophys. Res.* **91**, E218–E218 (1986).
7. Grimm, R. E. & Solomon, S. C. *Geophys. Res. Lett.* **14**, 538–541 (1987).
8. Turcotte, D. L. *J. geophys. Res.* (in the press).
9. Solomon, S. C. & Head, J. W. *J. geophys. Res.* **87**, 9236–9246 (1982).

10. Urey, H. C. in *Handbuch der Physik* Vol. 52, 363–418 (Springer, Berlin, 1959).
11. Mueller, R. F. *Icarus* **3**, 285–298 (1964).
12. Vinogradov, A. P. & Volkov, V. P. *Geochim. Int.* **8**, 463–467 (1971).
13. Lewis, J. S. *Earth planet. Sci. Lett.* **10**, 73–80 (1970).
14. Volkov, V. P., Zolotov, M. Yu. & Khodakovskiy, I. L. in *Chemistry and Physics of Terrestrial Planets* (ed. Saxena, S. K.) 136–190 (Springer, New York, 1986).
15. Barsukov, V. L., Surkov, Yu. A., Dmitriyev, L. V. & Khodakovskiy, I. L. *Geochim. Int.* **23**, 53–65 (1986).
16. Von Zahn, U., Kumar, S., Niemann, H. & Prinn, R. G. in *Venus* (eds Hunten, D. M., Colin, L., Donahue, T. M. & Moroz, V. I.) 299–430 (University of Arizona Press, Tucson, 1983).
17. Masursky, H. et al. *J. geophys. Res.* **85**, 8232–8260 (1980).
18. Seiff, A. in *Venus* (eds Hunten, D. M., Colin, L., Donahue, T. M. & Moroz, V. I.) 1045–1048 (University of Arizona Press, Tucson, 1983).
19. Marshall, J. R., Greeley, R., Tucker, D. W. & Pollack, J. B. *Icarus* **74**, 495–515 (1988).
20. Counsellman, C. C. III, Gourevitch, S. A., King, R. W., Lorient, G. B. & Prinn, R. G. *Science* **205**, 85–87 (1979).
21. Esposito, L. W., Knollenberg, R. G., Marov, M. Ya., Toon, O. B. & Turco, R. P. in *Venus* (eds Hunten, D. M., Colin, L., Donahue, T. M. & Moroz, V. I.) 474–564 (University of Arizona Press, Tucson, 1983).
22. Kite, F. T. & Wasson, J. T. *Science* **232**, 1225–1229 (1986).
23. Mason, B. (ed.) *Handbook of Elemental Abundances in Meteorites* (Gordon & Breach, New York 1971).
24. Basilevsky, A. T. & Head, J. W. A. *Rev. Earth planet. Sci.* **23**, 295–317 (1988).
25. Esposito, L. W. *Science* **223**, 1072–1074 (1984).
26. Esposito, L. W. *J. geophys. Res.* **93**, 5267–5276 (1988).
27. Parsons, B. *Geophys. J. R. ast. Soc.* **67**, 437–448 (1981).
28. Ronov, A. B. & Yaroshkevsky, A. A. *Geochim. Int.* **12**, 89–121 (1976).

Laboratory model of a planetary eastward jet

Joël Sommeria*, Steven D. Meyers & Harry L. Swinney

Center for Nonlinear Dynamics and the Department of Physics,
University of Texas, Austin, Texas 78712, USA

The concentration of eastward planetary currents into strong wavy jets is a major feature of the atmospheric general circulation (for example, the jet stream) and of oceanic currents (for example, the Gulf Stream east of Cape Hatteras). These strongly nonlinear jets are, however, not well understood. Here we describe experiments in which we have produced similar jets in a rotating annulus with mechanical forcing: the Coriolis force acting on fluid pumped radially inward from a ring of sources to a ring of sinks results in a strong eastward jet. For a wide range of parameters the jet has a robust wavy shape, as shown in Fig. 1. The wave velocity and wavelength of these Rossby waves are accounted for by simple scaling arguments. Experiments with dye injected into the fluid show that the jet acts as a barrier to tracer transport (see Fig. 1e), much as the southern polar night jet may act as a barrier to the transport of ozone from lower latitudes into the polar region¹. Our observations are insensitive to the details of the arrangement of the sinks and sources.

The annular tank (86.4 cm in diameter) is rotated rapidly (up to 4 Hz) to achieve an essentially two-dimensional geostrophic motion, as in large-scale planetary flows. (In a geostrophic flow the Coriolis force is balanced by pressure gradients².) The two-dimensionality of our flow is evident in photographs such as Fig. 1a, in which particles at different heights in the fluid are seen to produce parallel streaks. The rigidly rotating annulus has an inner radius $r_1 = 10.8$ cm, an outer radius $r_2 = 4r_1$, and a depth that increases linearly with radius from 17.1 cm at r_1 to 20.3 cm at r_2 . The conical bottom mimics the beta effect in a planetary atmosphere (the variation of the Coriolis force with latitude). The upper surface of the fluid (water) is in contact with a flat cover. In most of the experiments, including those shown in Fig. 1a, b and e, an eastward jet was produced by pumping fluid continuously from six sources in the bottom of the tank at $r = 35.1$ cm into six sinks at $r = 27.0$ cm; each source and sink consists of 400 holes (0.8 mm in diameter) within a 3.8-cm-diameter circle (these circles are inside the dark disks in Fig. 1b and e).

For low total pumping rates $F < 20 \text{ cm}^3 \text{ s}^{-1}$, each source or

sink produces a vortex that stays in the neighbourhood of a port, with sinks producing cyclonic vortices and sources producing anticyclonic vortices. For larger pumping rates, however, the vortices are advected eastward, resulting in a continuous jet. The strength of the jet is controlled by a balance of the total torque arising from the Coriolis force on the radially pumped fluid with the total torque arising from the friction force. The friction arises primarily from the Ekman boundary layers on the top and bottom surfaces of the tank; the tank was made deep to minimize the relative effect of this dissipation. The characteristic time for the dynamics, an oscillation period of the jet, is typically an order of magnitude shorter than the Ekman dissipation time $\tau = h_0/2(\Omega\nu)^{1/2}$ (where h_0 is the mean depth of the fluid, Ω is the angular velocity and ν is the kinematic viscosity); in other words, the Rossby number is typically 0.1. Because the dissipation is weak and the Coriolis force is large, the azimuthal velocity of the jet is typically two orders of magnitude greater than the average radial velocity, and the behaviour of the jet is determined mainly by inertial effects.

The waves on the jet are not attached to the sources and sinks—the crests move eastward with a speed of about 15% of the maximum velocity of the jet. The number of waves varies from 3 to 8, depending on Ω and F (the ranges studied are $6 < \Omega < 25 \text{ rad s}^{-1}$ and $15 < F < 370 \text{ cm}^3 \text{ s}^{-1}$). For a given Ω and F , the number of waves is reproducible but the wave amplitude and wavelength fluctuate in time. The quasi-steady wave pattern is the result of a balance of westward propagation of Rossby waves and eastward advection.

Quasi-geostrophic flows can be described in terms of the potential vorticity, $q(r, \theta)$, which is given in this situation (that is, uniform fluid density and a topography with a small constant slope) by²

$$q = \omega + \beta r \quad (1)$$

where $\omega = (\nabla \times \mathbf{u})_z$ is the vorticity of a fluid parcel relative to the rotating tank, $\beta = 2\Omega s/h_0$ is the coefficient of the beta effect ($s = -0.1$ is the slope of the bottom of the tank), r is the radial coordinate and $\mathbf{u}$ is the fluid velocity. In the absence of viscosity and forcing, q is conserved for each fluid parcel, that is,

$$Dq/Dt \equiv [\partial/\partial t + (\mathbf{u} \cdot \nabla)]q = 0 \quad (2)$$

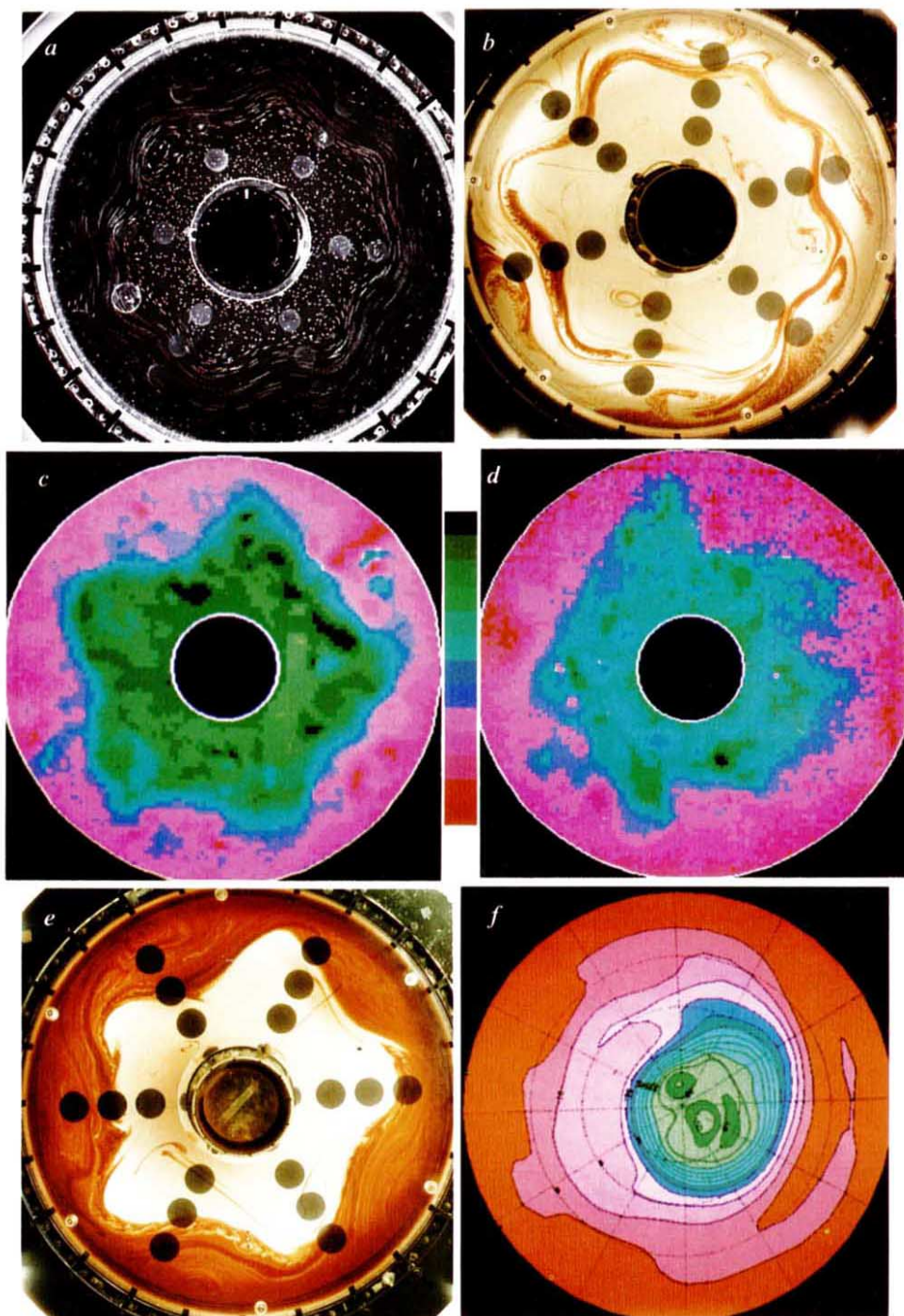
In our experiment the potential vorticity is nearly conserved—the weak dissipation is balanced on average by the forcing.

The velocity field in the annulus is determined from measurements of the length of particle streaks in time-exposure photographs, such as that in Fig. 1a. The velocity vectors from typically 300 particle streaks are fitted with cubic splines³, and the resultant velocity field is differentiated to determine the potential vorticity at any point in the fluid. We find that there is a steep gradient of potential vorticity in the middle of the jet, as Fig. 1c and d illustrates, whereas away from the centre of the jet the potential vorticity is fairly uniform. (A gradient in potential vorticity is necessary for the propagation of Rossby waves.) Juckes and McIntyre⁴ observed a similar steepening of an initially smooth gradient in their numerical study of a single-layer model of a terrestrial stratospheric polar vortex. At low latitudes the initially axisymmetric flow around the vortex was unstable and rapidly became turbulent, creating a well mixed state, whereas the polar vortex remained an essentially isolated structure. Mixing between these two regions took the form of long thin 'tongues' of potential vorticity and was associated with stratospheric wave-breaking^{4–6}. This local steepening of potential vorticity gradients seems to be common behaviour for strongly stirred inertial quasi-geostrophic layers and is similar to the steepening of density gradients observed in strongly stirred stratified non-rotating fluids⁷.

In our experiment, injected dye is focused into the jet and renders it visible even when the dye is injected at a radius only occasionally reached by the wave crests or troughs. In a two-dimensional non-divergent flow this convergence must be associ-

* Permanent address: Madylam, ENSHMG, BP 95, 38402 St Martin D'Hères Cedex, France.

Fig. 1 *a*, A jet with eight waves, visualized with particle streaks photographed in a frame of reference rotating with the tank ($\Omega = 25.1 \text{ rad s}^{-1}$, $F = 30 \text{ cm}^3 \text{ s}^{-1}$; exposure time = 0.5 s). *b*, A jet with seven waves, visualized by continuous injection of dye through a small tube ($\Omega = 18.8 \text{ rad s}^{-1}$, $F = 40 \text{ cm}^3 \text{ s}^{-1}$). *c*, Potential vorticity q for a jet with five waves. The potential vorticity, deduced from particle streak measurements, ranges from -8.1 rad s^{-1} (red) to -0.6 rad s^{-1} (dark green) ($\Omega = 18.8 \text{ rad s}^{-1}$, $F = 90 \text{ cm}^3 \text{ s}^{-1}$). *d*, Potential vorticity q for a jet that formed spontaneously in a slowly decelerating tank with the sources and sinks alternating in a ring at $r = 27.0 \text{ cm}$. The potential vorticity ranges from -15.0 rad s^{-1} (red) to 2.3 rad s^{-1} (dark green) ($\Omega = 25.1 \text{ rad s}^{-1}$, decreasing at a rate of 0.012 rad s^{-2} ; $F = 137 \text{ cm}^3 \text{ s}^{-1}$). *e*, A jet with five waves, photographed more than 500 revolutions of the tank after dye was injected outside the jet ($\Omega = 18.8 \text{ rad s}^{-1}$, $F = 90 \text{ cm}^3 \text{ s}^{-1}$). In *a*, *b* and *e* the jet was produced by pumping fluid from an outer ring of sources to an inner ring of sinks. *f*, Ertel's potential vorticity on an isentropic surface in the upper stratosphere on a calm winter day in the Southern Hemisphere (from ref. 27); contour lines at intervals of $5 \times 10^{-8} \text{ Pa}^{-1} \text{ s}^{-1}$ are distinguished by colour. (The units are those of potential vorticity appropriate to the atmosphere.) A strong potential vorticity gradient in the central region of the jet isolates the polar vortex from the surrounding atmosphere.



ated with an increase of the jet path length, resulting in an increase of the wave amplitude which cannot persist in a steady regime. The growth is limited by a random localized ejection of fluid from the jet which can be observed in the motion of particles suspended in the flow. A particle typically stays inside the jet for several revolutions and is then advected far outward at a wave crest by one of the weak recirculating eddies surrounding the jet. Particles are also advected inwards at wave troughs, but less often. The ejection can be seen in Fig. 1*b* and *e* as tongues of dye. This process slowly fills the space outside the jet with a uniform concentration of dye. When dye is injected far from the jet, it is spread uniformly in the region delimited by the jet, but there is remarkably little transport across the jet for very long periods of time, as Fig. 1*e* illustrates.

These diffusion properties elucidate the dynamics of the potential vorticity, which is advected like a dye; compare the

photograph in Fig. 1*e* of a five-wave jet visualized with dye with the plot in Fig. 1*c* of the potential vorticity determined for a five-wave jet. The flow toward the centre of the jet brings the positive potential vorticity produced by the sinks close to the negative potential vorticity produced by the sources, thus inducing a steep gradient. The local ejection of potential vorticity associated with the ejection of fluid particles into the surrounding regions helps to mix potential vorticity and contributes to the small-scale structure seen in Fig. 1*e*. In regions of quasi-uniform potential vorticity the dye spreads quickly and evenly but does not cross the strong potential vorticity gradient in the middle of the jet. The zones of uniform potential vorticity, which cannot support Rossby waves, confine the waves by a nonlinear reflection effect^{8,9}. The ejection of particles at wave crests is probably related to the filamentation of potential vorticity observed by Jukes and McIntyre⁴, as mentioned above. The

filamentation process has been also observed recently by Dritschel¹⁰ in near-inviscid numerical computations for a circular blob of uniform vorticity surrounded by irrotational fluid. He found that a small perturbation on an initially symmetric vortex creates an irregular region of ever-increasing complexity around the vortex.

Our results are insensitive to the details of the forcing geometry, provided that the flow is eastward with respect to the forcing ports. For example, closing all of the ports on one side of the tank makes virtually no difference to the behaviour of the flow if the total flux F is the same—dye injected away from the centre of the jet is still inhibited from crossing the jet, even in the unforced part of the annulus. Also, a narrow coherent wavy jet is still observed when the distance between the sources and sinks is doubled. A similar, although somewhat less regular, structure is obtained even using a single ring of alternating sources and sinks together with a mean eastward flow produced by a slow deceleration of the tank; see Fig. 1*d*. Seen in a different frame of reference, this situation is equivalent to a forcing that propagates westward in a fluid at rest; this forcing has been used by others^{11–13} to demonstrate the production of a mean zonal flow by eddies¹⁴. In our experiment, in which the effect of friction is less important than in previous experiments, the propagating forcing leads to an inertial eastward jet.

The total angular momentum of the jet results from a balance of the total torque introduced by the forcing with the total torque caused by friction in the boundary layers. This balance leads to the relation $UL = (\Omega \tau d / \pi \langle r \rangle h_0) F$, where the distance d between the rings of sources and sinks is assumed to be small compared with the mean radius $\langle r \rangle$ of these two rings. Here U is a velocity scale and L is the characteristic jet width. Considering the balance between advection and the beta effect yields the classic condition $U/L^2 \approx \beta$. The two relations involving U and L lead to¹⁵

$$U = a(sd^2/2\pi^2\nu h_0\langle r \rangle^2)^{1/3}(F\Omega)^{2/3} \quad (3)$$

$$L = b(h_0d/4\pi s\langle r \rangle\nu^{1/2})^{1/3}\Omega^{-1/6}F^{1/3} \quad (4)$$

where a and b are non-dimensional parameters that in general depend on the velocity profile and the exact definitions of U and L . The observed azimuthally averaged velocity, $U(r) = \langle u_\theta(r, \theta) \rangle_\theta$, fits a $\text{sech}^2(r/L)$ profile fairly well for the full range of experimental conditions examined. Moreover, the data are described reasonably well by the scaling relations (3) and (4) with a and b given by the values for a sech^2 profile, $a = (3/8)^{1/3}$ and $b = (1/3)^{1/3}$, as Fig. 2*a* shows for the velocity data. Thus the jet has the same basic structure over a wide range of parameters.

Some qualitative aspects of the jet can be understood using linear theory. Because the perturbations are large, the best form for modelling the zonal flow is not obvious; for simplicity we take the azimuthally averaged flow. The measured profile satisfies the Rayleigh-Kuo criterion of marginal stability¹⁶, which states that the maximum of $d^2\langle U(y) \rangle / dy^2$ is equal to $|\beta|$ (where the jet curvature is neglected and y represents the radial direction). When the forcing is turned on, the initially unstable jet profile is broadened by the growth of instabilities until it is stabilized by the beta effect into a marginally stable shape. There are two predicted modes of linear instability, varicose and sinuous, but only the latter is unstable for an eastward jet satisfying the Rayleigh-Kuo marginal stability condition¹⁷; this explains our observation of a wavy shape. The wavenumber k for the instability is $2^{1/2}L^{-1}$ for a jet with a sech^2 profile. The observed number of waves is in qualitative but not quantitative agreement with this linear theory, as Fig. 2*b* shows. The departure from linear theory is not surprising as the experiments are in a nonlinear regime involving finite amplitude waves and wave-breaking.

Our flow is barotropic (the density is uniform and the velocity field does not vary with height), whereas the atmosphere and

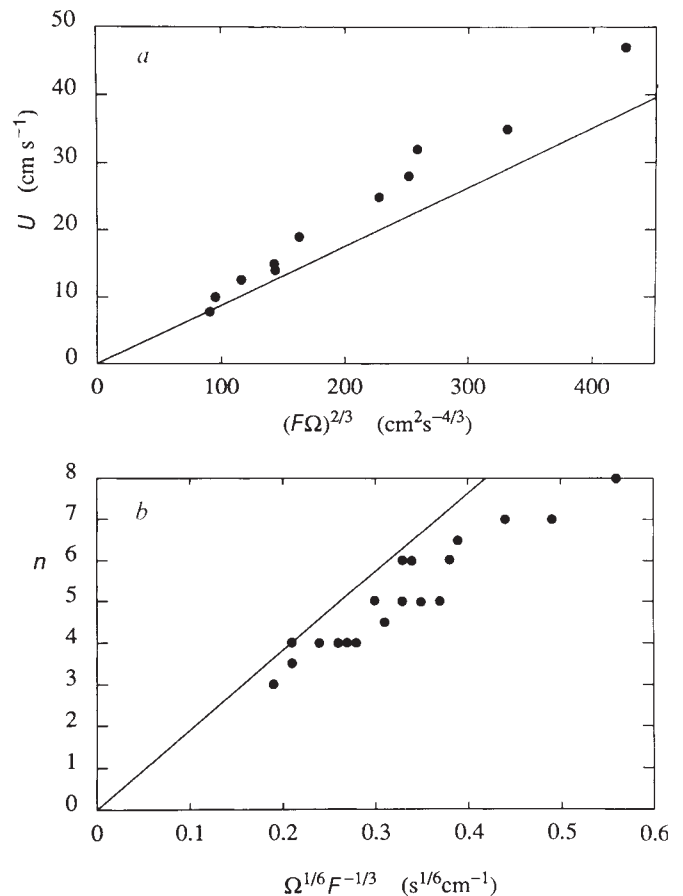


Fig. 2 *a*, The maximum of the azimuthally averaged velocity of the jet, U , compared with the prediction of a marginal stability analysis (diagonal line). *b*, The observed number of waves compared with the value predicted by linear theory (solid line), $n = k(r) = 2^{1/2}L^{-1}\langle r \rangle$. In the neighbourhood of a transition the number of waves fluctuates between n and $n+1$; these points are plotted as half-integer values of n .

oceans are baroclinic (the density and velocity depend on altitude), although planetary-scale flows in the middle stratosphere often exhibit deep vertical structure similar to barotropic flows. Some elegant laboratory studies of Rossby waves have been conducted previously on a baroclinic system, a heated rigidly rotating annulus of water^{18–22}. Wavy coherent jets were obtained over a wide range of parameters, but their three-dimensional baroclinic structure is difficult to analyse. We chose to study a barotropic system because it is simpler, yet it exhibits many of the phenomena exhibited by baroclinic flows, including, in particular, Rossby waves. The observation that phenomena found in barotropic flows also occur in baroclinic flows has been termed the 'principle of barotropic analogy' by McWilliams²³.

The behaviour of the potential vorticity in our experiments is analogous to the behaviour of Ertel's potential vorticity in the atmosphere on an isentropic surface²⁴. Above the lower troposphere in the Earth's atmosphere, where most of the diabatic and friction effects occur, the flow is dominated by advective effects and Rossby wave propagation; gradients of the potential vorticity on an isentropic surface are due to the tilt of the constant-density surfaces and the curvature of the Earth. Rossby waves propagating westward in an upper atmospheric layer or in the ocean are often amplified by the baroclinic instability²⁴, just as a steady Rossby wave is amplified by the forcing in our tank. The concentration of the potential vorticity gradient in our jet surrounded by zones of uniform potential

vorticity is similar to that observed in the atmosphere and the ocean^{25,26}; see, for example, Fig. 1*f*. Filamentary mixing across a potential vorticity gradient, such as that illustrated by Fig. 1*e* and *f*, may also play an important role in mixing of atmospheric tracer elements such as ozone^{1,27,28}.

In conclusion, we have described a relatively simple system for studies of instabilities and turbulence in barotropic jets at high Reynolds number ($\sim 10^5$). Our experiments suggest that jets in planetary scale flows can be inertial structures that are largely independent of the details of the forcing, depending mainly on the total energy and momentum. We find that the wavy eastward jets serve as a striking barrier to particle transport. The potential vorticity barrier has also been found to be remarkably complete in numerical experiments⁴. The laboratory and numerical studies and an analysis of data on the Antarctic atmosphere have led to the vortex-isolation hypothesis of McIntyre²⁹: "in the real Antarctic lower stratosphere, the sealing-off of the vortex core from its surroundings might be practically complete".

The narrow wavy structure found for eastward jets in our experiments contrasts markedly with the turbulent flow containing persistent vortices found for westward jets in earlier experiments in the same apparatus¹⁵. The relevance of these observations to planetary-scale flows should be examined further in future studies.

This research is supported by the Office of Naval Research Nonlinear Dynamics Program, the National Science Foundation

Program in Fluid Mechanics and Hydraulics and NATO Grant for International Collaboration. One of us (J.S.) is supported in part by the Centre National de la Recherche Scientifique.

Received 13 July 1987; accepted 3 November 1988.

1. Leovy, C. B. *et al. J. Atmos. Sci.* **42**, 230–244 (1985).
2. Pedlosky, J. *Geophys. Fluid Dyn.* (Springer, New York, 1982).
3. Nguyen Duc, T. & Sommeria, J. *J. Fluid Mech.* **192**, 175–192 (1988).
4. Jukes, M. N. & McIntyre, M. E. *Nature* **328**, 590–596 (1987).
5. McIntyre, M. E. & Palmer, T. N. *Pure appl. Geophys.* **123**, 964–975 (1985).
6. Rood, R. B. *Pure appl. Geophys.* **123**, 733–755 (1985).
7. Turner, J. S. *Buoyancy Effects in Fluids* Ch.9.1.1 (Cambridge Univ. Press, 1973).
8. Talley, L. D. *J. phys. Oceanogr.* **13**, 972–987 (1983).
9. Killworth, P. D. & McIntyre, M. E. *J. Fluid Mech.* **161**, 449–492 (1985).
10. Dritschel, D. G. *J. Fluid Mech.* **194**, 511–547 (1988).
11. Colin de Verdiere, A. *J. Fluid Mech.* **94**, 39–64 (1979).
12. McEwan, A. D., Thomson, R. O. R. Y. & Plumb, R. A. *J. Fluid Mech.* **99**, 655–672 (1980).
13. Whitehead, J. A. *Tellus* **27**, 358–363 (1975).
14. Rhines, P. B. & Holland, W. R. *Dyn. Atmos. Oceans* **3**, 289–325 (1979).
15. Sommeria, J., Meyers, S. D. & Swinney, H. L. *Nature* **331**, 689–693 (1988).
16. Kuo, H. L. *J. Met.* **6**, 105–122 (1949).
17. Howard, P. G. & Drazin, L. N. *J. math. Phys.* **43**, 83–99 (1964).
18. Fultz, D. *et al. Met. Monogr.* **4**, 1–104 (1959).
19. Pfeffer, R. L. & Fowles, W. W. *J. Atmos. Sci.* **25**, 361–371 (1968).
20. Hide, R. & Mason, P. J. *Adv. Phys.* **24**, 47–100 (1975).
21. Hart, J. E. A. *Rev. Fluid Mech.* **11**, 147–172 (1979).
22. White, H. D. & Koschmeider, E. L. *Geophys. Astrophys. Fluid Dyn.* **18**, 301–310 (1981).
23. McWilliams, J. in *Non-Linear Topics in Ocean Physics* (ed. Osborne, A.) (North-Holland, Amsterdam, in the press).
24. Hoskins, B. J., McIntyre, M. E. & Robertson, A. W. *Q.J.R. met. Soc.* **111**, 877–946 (1985).
25. Platzman, G. W. *Tellus* **1**, 53–64 (1949).
26. McDowell, S., Rhines, P. & Kefler, T. *J. phys. Oceanogr.* **12**, 1417–1436 (1982).
27. Al-Ajmi, D. N., Harwood, R. S. & Miles, T. *Q.J.R. met. Soc.* **111**, 359–389 (1985).
28. Butchart, N. & Remsburg, E. E. *J. Atmos. Sci.* **43**, 1319–1339 (1986).
29. McIntyre, M. E. *J. Atmos. terr. Phys.* (submitted).

Geochemical evidence for suppression of pelagic marine productivity at the Cretaceous/Tertiary boundary

James C. Zachos*, Michael A. Arthur* & Walter E. Dean†

* Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island 02882, USA

† US Geological Survey, PO Box 25046, Denver, Colorado 80225, USA

The normal, biologically productive ocean is characterized by a gradient of the $^{13}\text{C}/^{12}\text{C}$ ratio from surface to deep waters. Here we present stable isotope data from planktonic and benthic microfossils across the Cretaceous/Tertiary boundary in the North Pacific, which reveal a rapid and complete breakdown in this biologically mediated gradient. The fluxes of barium (a proxy for organic carbon) and CaCO_3 also decrease significantly at the time of the major marine plankton extinctions. The implied substantial reduction in oceanic primary productivity persisted for ~ 0.5 Myr before the carbon isotope gradient was gradually re-established. In addition, the stable isotope and preservational data indicate that environmental change, including cooling, began at least 200 kyr before the Cretaceous/Tertiary boundary, and a peak warming of $\sim 3^\circ\text{C}$ occurred 600 kyr after the boundary event.

The section at DSDP Site 577 on the Shatsky Rise ($32^\circ 26' \text{N}$, $157^\circ 43' \text{E}$) was collected by hydraulic piston core and yielded the first continuous, undisturbed sequence of pure nannofossil ooze ($>90\%$ CaCO_3) across the Cretaceous/Tertiary (K/T) boundary. Because of the relatively shallow burial depth of the K/T boundary sequence ($<110\text{m}$ below the sea floor (m.b.s.f.)¹), the calcareous microfossils from Site 577 have not undergone significant recrystallization related to burial^{2–4}. The position of the K/T boundary is placed at the last appearance of planktonic foraminifera of the *Abathomphalus mayaroensis* zone⁴ and first appearance of calcareous nannofossils of the *Marcalius inversus* zone², which also coincides with peak abun-

dance of magnetite–glauconite spherules at 109.09 m.b.s.f. (refs 4, 5). A peak irridium anomaly of 61 ng cm^{-2} also occurs at the boundary, as defined by nannofossils, in adjacent Hole 577B (ref. 6). The palaeodepth of Site 577 is estimated as having been 2,400 m at the time of the K/T boundary³. The record at DSDP Site 577 is adequate to resolve the nature of environmental changes in pelagic settings that occurred in conjunction with the K/T boundary event, and provides important constraints on hypotheses for extinction.

To reconstruct changes in the surface-to-deep-water carbon isotope gradient of total dissolved carbon (TDC), stable-isotope analyses were conducted on fine fraction ($<63\text{-}\mu\text{m}$) carbonate and several species of planktonic and benthic foraminifera in closely spaced samples over a 10-m interval spanning the K/T boundary^{3,4,7} (Fig. 1). The $\delta^{13}\text{C}$ of fine-fraction carbonate (primarily calcareous nannoplankton) and of planktonic foraminifera from this site decreases by about 1.5‰ at the K/T boundary relative to values just below the boundary. The $\delta^{13}\text{C}$ values of biserial benthic foraminifera *Aragonia* and of trochospiral benthic foraminifera *Nuttallides*, *Gavelinella* and *Gyroidinoides* show an enrichment across the boundary converging with $\delta^{13}\text{C}$ values of both planktonic foraminifera and fine-fraction carbonate just above the boundary (Fig. 2).

Convergence of $\delta^{13}\text{C}$ values of planktonic foraminifera, nannoplankton (fine-fraction) and benthic foraminifera indicates that the biologically mediated gradient in carbon isotopes between the surface and deep waters, which is characteristic of productive oceans, effectively disappeared at the K/T boundary. Such conditions have been termed a 'Strangelove' ocean on the basis of predictions for events following a mass extinction^{8,9}. The loss of a surface-to-deep-water carbon isotope gradient could only be caused by a major reduction in primary production or by extremely rapid rates of oceanic turnover. At Site 577 the lack of a carbon isotope gradient suggests that a poorly productive ocean persisted for at least 0.5×10^6 yr following the K/T boundary event (Fig. 2). The gradual development of a surface-to-deep-water $\delta^{13}\text{C}$ gradient in the period following this interval probably represents a progressive increase in primary production.

The changing offset of $\delta^{13}\text{C}$ between biserial and trochospiral

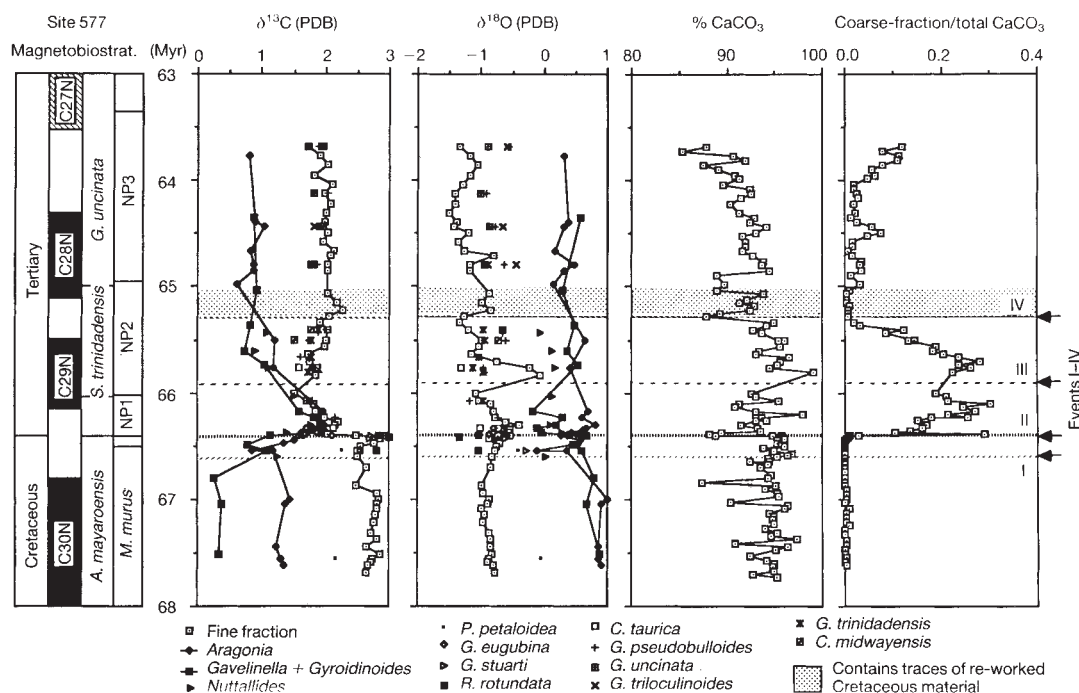


Fig. 1 Age plots of carbon and oxygen stable isotope compositions of fine-fraction carbonate ($<63 \mu\text{m}$), benthic foraminifer *Aragonia*, trochospiral genera *Gavelinella* + *Gyroidinoides*, and *Nuttallides*, various species of planktonic foraminifera, and carbonate content of whole rock samples. All stable isotope data are expressed in the delta notation, δ , in ‰ relative to the PDB marine carbonate standard²⁷. The final column shows the weight ratio (%) of coarse fraction ($>63 \mu\text{m}$) to total carbonate. Age determinations were made using available magnetostratigraphy²⁸ and nanofossil stratigraphy¹³ (presented on the far left) and the most recently published timescale²⁹. Four separate events are recognized in the evolution of the stable isotope and carbonate content records and are denoted by the four dashed lines. The shaded area represents a portion of the sequence which contains re-worked Cretaceous fossils. The gap in sampling from 65.8 to 65.95 Myr encompasses the interval of an organic geochemistry core removed on board ship.

benthic foraminiferal genera (Fig. 1) is particularly interesting because of the observation that many species of modern benthic foraminifera do not reside directly on the sea floor but instead within the upper few centimetres of sediment, and therefore secrete their tests in carbon isotope equilibrium with the surrounding pore waters rather than with bottom waters¹⁰. *In situ* decay of organic matter in the upper few centimetres of sediment produces a pore-water $\delta^{13}\text{C}$ gradient with a magnitude which is, in part, a function of the accumulation rate of organic carbon (C_{org})¹¹. Convergence of $\delta^{13}\text{C}$ values of *Aragonia* and that of *Gavelinella* + *Gyroidinoides* across the K/T boundary may indicate that the pore-water $\delta^{13}\text{C}$ gradient is eliminated as a result of the reduction in one C_{org} flux to sediments. The absence of benthic foraminiferal carbon isotope differences during the recovery period suggests that new C_{org} flux levels were lower than that of the latest Cretaceous.

Other evidence which suggest earliest Palaeocene reduced primary productivity involves carbonate and barium mass-accumulation-rate data. In the absence of changes in the rate of dissolution, the rate of accumulation of biogenic carbonate should be proportional to production in surface water. Carbonate accumulation rates decline on average by a factor of four across the K/T boundary in all pelagic marine sequences examined³. In general, preservation of calcareous microfossils improves across the K/T boundary at most localities^{4,12}. Therefore, the decrease in carbonate accumulation rates from Maestrichtian to the early Palaeocene cannot be easily attributed to increased rates of dissolution and must have resulted from a decrease in the rate of CaCO_3 production in surface waters.

The distribution of dissolved barium in the modern oceans is nutrient-like and appears to be mediated by biological activity^{13,14}. In the modern Pacific Ocean, Ba and organic-carbon sediment-trap fluxes are commonly proportional¹⁵. Because

sediment pore waters are near saturation with respect to barite¹⁶, barite is typically preserved in sediments whereas C_{org} is oxidized. Thus, Ba accumulation rates can be used as a proxy for C_{org} flux to the sediments¹⁷. At Site 577, Ba accumulation rates decrease from >0.2 to $<0.1 \text{ mg cm}^{-2} \text{ per } 10^3 \text{ yr}$ across the K/T boundary (Fig. 3), reflecting the decreased rate of C_{org} .

We have recognized four discrete episodes in the evolution of the stable-isotope and carbonate records across the K/T boundary and in the early Palaeocene at Site 577. The first of these events begins at least 10^5 yr before the main plankton extinctions and the last occurs nearly $1.5 \times 10^6 \text{ yr}$ later. We believe that these events represent significant changes in oceanic fertility, circulation and climate, which are accompanied by consequent effects on biotic evolution.

Event I (66.6 Myr), which precedes the K/T boundary event, is marked by enrichment in both planktonic and benthic microfossil carbon isotope records, accompanied by a slight increase in carbonate content and an improvement in microfossil preservation. In addition, Ba accumulation rates decrease. The increase in $\delta^{13}\text{C}$ of both the planktonic and benthic records probably indicates a change in $\delta^{13}\text{C}$ of the oceanic TDC reservoir. This could have resulted from increased organic-carbon burial on continental margins (or decreased carbon fluxes from weathering) that are related to the latest Cretaceous transgression, which began at about 66.7 Myr (ref. 18). In addition, there may have been changes in deep-water sources, increased oxygenation of deep waters and/or lower organic-carbon flux at Site 577 at this time. Enriched fine-fraction $\delta^{18}\text{O}$ values also indicate cooling of surface waters; this suggests that the rate of oceanic turnover and ventilation of the deep reservoir increased before the K/T boundary as the result of global cooling and intensification of thermohaline-driven circulation. The increase in relative proportions of the temperate planktonic foraminiferal

Fig. 2 Expanded 2-Myr record across the K/T boundary of the stable carbon and oxygen isotope data from Fig. 1. Shaded area represents a zone containing traces of re-worked Cretaceous microfossils.

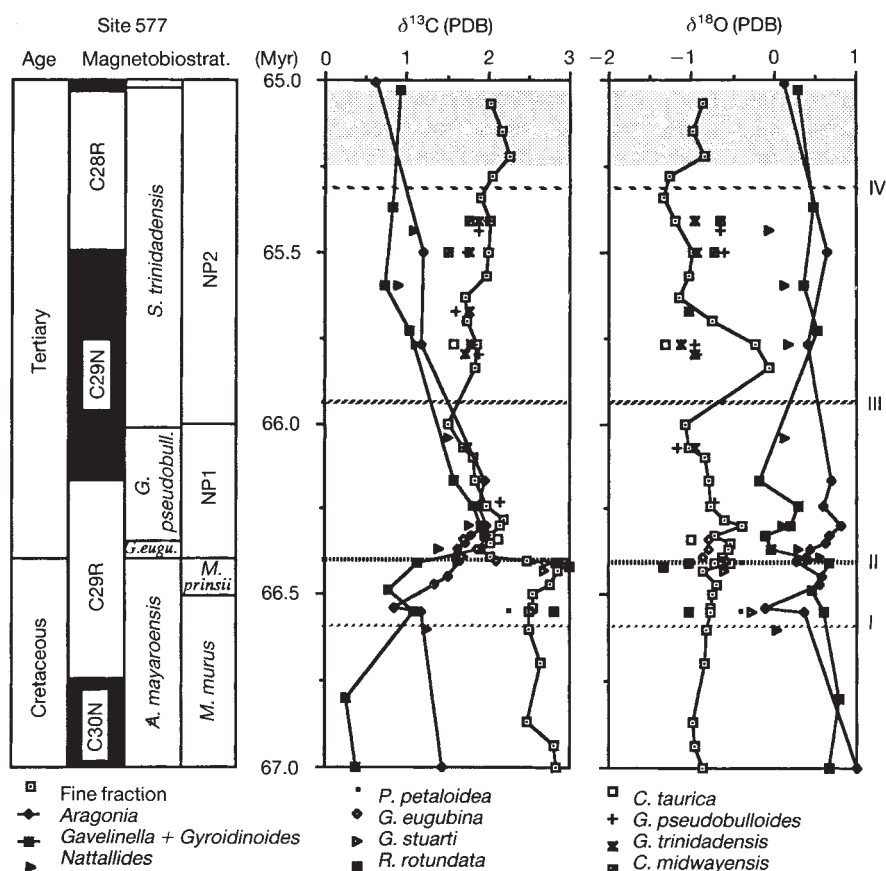
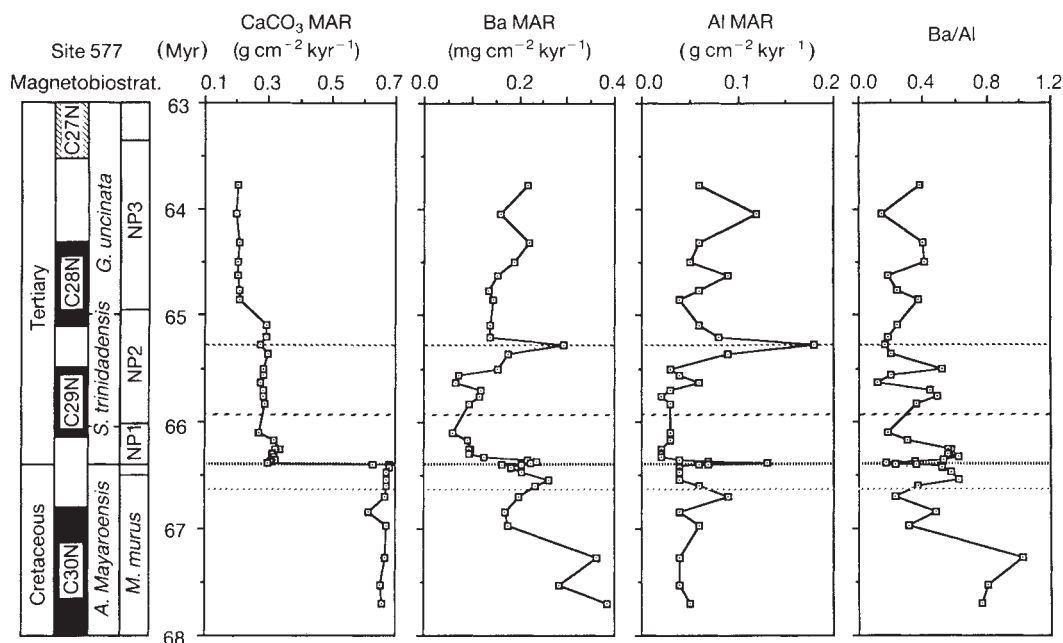


Fig. 3 Age plots of barium (Ba), aluminium (Al) and carbonate mass accumulation rates (MAR), and Ba/Al ratios, at Site 577. The accumulation rates were calculated using physical properties data from the DSDP Leg 86 volume and sedimentation rates based on the magneto- and nannofossil stratigraphy at left. Age assignments for the immediate post-boundary interval were averaged from the top of the Cretaceous (109.1 m.b.s.f., 66.4 Myr) to the base of Chron 29N (108.4 m.b.s.f., 66.17 Myr). As a result, we cannot resolve changes in accumulation rates that occurred on timescales of 10^4 yr or less at Site 577. Because the Base of Chron 30N (C30N) or the *M. murus* zone were not reached at Hole 577, the thickness of the *M. murus* zone (8.48 m) was extrapolated from adjacent Hole 577A. The Ba/Al ratio represents weight ratio of Ba to Al_2O_3 .



genus *Heterohelicidae* at this low latitude provides additional support for cooling before the K/T boundary event⁴.

Event II: K/T boundary (66.4 Myr), the main plankton extinction and productivity crisis, features the rapid collapse of the surface-to-deep-water carbon isotope gradient and decreases in carbonate and Ba accumulation rates described above. A puzzling aspect of this event is the 100-fold increase recorded in the concentration of foraminifera relative to total carbonate

(Fig. 1). There are three possible explanations for this phenomenon: (1) intensified deep circulation resulting in winnowing of the fine fraction; (2) less fragmentation and improved preservation of the more dissolution-susceptible planktonic foraminifera as a result of deepening of the carbonate compensation depth (CCD) and/or lowered rates of *in situ* dissolution brought about by decreased C_{org} decay within sediment pore waters; and (3) an ecological trophic-structure response to the

extinctions and environmental change in which calcareous nanoplankton are replaced by siliceous or organic-walled phytoplankton. We favour preservation changes (2) as the possible origin of the increase in foraminifera concentrations. It is generally accepted that coccoliths tend to be more dissolution-resistant than foraminifera and that calcite dissolution above the calcite saturation horizon is driven mainly by titration by metabolic CO_2 derived from C_{org} decay at or near the sediment-water interface¹⁹. Dissolution at the shallow palaeodepth of the Shatsky Rise near site 577 would have been controlled to a large extent by metabolic CO_2 production. Decay of C_{org} at the sediment-water interface would decline as a result of reduced C_{org} flux during the 'Strangelove' ocean event making bottom and pore water less corrosive to calcite. Furthermore, taking into account the large initial decrease in CaCO_3 accumulation that occurred as a result of the productivity crisis, and assuming that riverine input of inorganic C and alkalinity remained constant, one would expect a rapid deepening of the CCD and much improved preservation of foraminifera across the K/T boundary. Thus we explain the increased foraminiferal flux in the Palaeocene, implied by the higher relative proportion of foraminifers, by enhanced preservation.

Event III (65.9 Myr) is characterized by the gradual return of the surface-to-deep-water carbon isotope gradient, a slight increase in Ba accumulation rates and a gradual decrease in the planktonic foraminifera concentrations. These changes indicate that surface-water productivity began to increase ~ 0.5 Myr after the K/T boundary event, albeit at relatively low levels. Climate fluctuations of several degrees are indicated by variations in nannofossil and foraminiferal $\delta^{18}\text{O}$. Rates of evolution of planktonic forams remained high²⁰, suggesting continuing environmental instability.

Event IV (65.3 Myr) signals the full expression of the surface-to-deep-water carbon isotope gradient and a decrease in planktonic foraminiferal concentrations to normal values. Aluminium accumulation rates also increased at this time. Because Al accumulation rates are probably a reflection of flux of aeolian detritus to this site²¹, the increased Al accumulation rates may represent climate change at this time and possibly an increase in wind strength. Environmental stability is indicated by a nearly constant $\delta^{13}\text{C}$ gradient, relatively stable $\delta^{18}\text{O}$ values and low rates of change of planktonic foraminiferal faunas in the *G. uncinata* zone²⁰.

The extinction of phyto- and zooplankton and disappearance of the carbon isotope gradient is abrupt and occurs over a period of less than 10^4 yr. It is unlikely that a gradual environmental change occurring over a longer period of time (that is, 10^5 – 10^6 yr) could have resulted in such a rapid and extreme collapse of the marine ecosystem unless some as yet unknown environmental threshold was exceeded²². The instantaneous environmental stress that would be brought about by a bolide impact provides a possible explanation for the collapse of primary productivity and its coincidence with an abundance of spherules and platinum-group metals at the K/T boundary, as exhibited at Site 577. It fails to explain, however, why primary productivity remained low for such a long period after the event.

It is possible that the 0.5-Myr 'Strangelove' ocean may represent the normal recovery time for the biosphere following a sudden mass extinction event. An alternative explanation is that the slow recovery of primary productivity was the result of climate and oceanic circulation instability. The rate of speciation was extremely high for the first several 10^6 years above the K/T boundary at Site 577, with one assemblage of planktonic foraminifera quickly replacing another²⁰. Because the distributions of planktonic foraminifera are limited to particular water masses, the high turnover rate of assemblages suggests that surface-water conditions changed rapidly during the Danian. As a consequence, an efficient, highly structured ecosystem may have been incapable of developing. The origin of this climatic instability is uncertain but may prove to be related to feedback

mechanisms involving the global carbon cycle^{9,23–25}. Models predict that, following a rapid extinction event, atmospheric p_{CO_2} would have increased by as much as a factor of two or three on a timescale of several oceanic mixing cycles²⁶, as the result of nearly complete cessation of oceanic productivity. Such increases in p_{CO_2} would be expected to result in global warming. The nannofossil data, which represent the most complete time series, suggest a gradual 3°C warming of surface waters during the initial 0.5 Myr of the Palaeocene.

We thank D. McLean for comments on this manuscript, and M. Bender, K. Caldeira, B. Corliss, D. Fastovsky, A. Mackensen, D. McCorkle, K. Miller, C. Officer, M. Rampino and R. Thunell for discussions. Special thanks are extended to J. Gerstel for assistance during sample preparation. This research was supported by the NSF.

Received 11 April; accepted 9 November 1988.

1. Heath, G. R. *et al. Init. Rep. DSDP* **86** (1985).
2. Monechi, S. *Init. Rep. DSDP* **86**, 301–306 (1985).
3. Zachos, J. C. & Arthur, M. A. *Paleoceanography* **1**, 5–26 (1986).
4. Gerstel, J., Thunell, R. C., Zachos, J. C. & Arthur, M. A. *Paleoceanography* **1**, 97–117 (1986).
5. Smit, J. & Romein, A. J. T. *Earth planet. Sci. Lett.* **74**, 155–170 (1985).
6. Michel, H. V., Asaro, F., Alvarez, W. & Alvarez, L. W. *Init. Rep. DSDP* **86**, 533–538 (1985).
7. Zachos, J. C., Arthur, M. A., Thunell, R. C., Williams, D. F. & Tappa, E. J. *Init. Rep. DSDP* **86**, 513–532 (1985).
8. Broecker, W. S. & Peng, T. H. *Tracers in the Sea* (ELDIGIO, New York, 1982).
9. Hsu, K. J. & McKenzie, J. A. in *The Carbon Cycle & Atmospheric CO_2 : Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 487–492 (Am. Geophys. Un., Washington DC, 1985).
10. Corliss, B. H. *Nature* **314**, 435–438 (1985).
11. McCorkle, D. C., Emerson, S. R. & Quay, P. D. *Earth planet. Sci. Lett.* **74**, 13–26 (1985).
12. Thierstein, H. R. *Soc. Econ. Paleontol. Miner. spec. Publ.* **32**, 355–394 (1981).
13. Dehairs, F., Lambert, C. E., Chesselet, R. & Risler, N. *Biogeochemistry* **4**, 119–139 (1987).
14. Chan, L. H., Drummond, D., Edmond, J. M. & Grant, B. *Deep Sea Res.* **24**, 613–649 (1977).
15. Dymond, J. *Am. Geophys. Un. Ocean Sci. Mtg* (1986).
16. Church, T. M. in *Reviews in Mineralogy*, Vol. 6 (ed. Burns, R. G.) 175–209 (Miner. Soc. Am., Washington DC, 1976).
17. Schmitz, B. *Paleoceanography* **2**, 63–78 (1987).
18. Haq, B. U., Hardenbol, J. & Vail, P. R. *Science* **235**, 1156–1167 (1987).
19. Emerson, S. E., Bender, M. A. *J. Mar. Res.* **39**, 139–162 (1981).
20. Gerstel, J., Thunell, R. & Ehrlich, R. *Geology* **15**, 665–668 (1987).
21. Doh, S.-J., King, J. W. & Leinen, M. *Paleoceanography* **3**, 89–112 (1988).
22. Crowley, T. S. & North, G. R. *Science* **240**, 996–1002 (1988).
23. McLean, D. M. *Cret. Res.* **6**, 235–259 (1985).
24. Kasting, J. F., Richardson, S. M., Pollack, J. B. & Toon, O. B. *Am. J. Sci.* **286**, 361–389 (1986).
25. Rampino, M. & Volk, T. *Nature* **332**, 63–65 (1988).
26. Caldeira, K. G., Rampino, M. R. & Zachos, J. C. *Eos* **69**, 377 (1988).
27. Craig, H. *Geochim. cosmochim. Acta* **12**, 133–149 (1955).
28. Bleil, U. *Init. Rep. DSDP* **86**, 441–458 (1985).
29. Berggren, W. A., Kent, D. V., Flynn, J. J. & Van Couvering, J. A. *Bull. geol. Soc. Am.* **96**, 1407–1418 (1985).

A brachiopod calcite record of the oceanic carbon and oxygen isotope shifts at the Permian/Triassic transition

Michał G. Szczyński, Stanisław Hałas*, Antoni Hoffman & Krzysztof Małkowski

Institute of Paleobiology, Polish Academy of Sciences, Al. Żwirki i Wigury 93, PL-02-089 Warszawa, Poland

* Institute of Physics, Maria Curie-Skłodowska University, Pl. Marii Skłodowskiej-Curie 1, PL-20-031 Lublin, Poland

The Permian/Triassic (P/T) mass extinction was the most dramatic of all such events but its cause has thus far remained elusive^{1–3}. Carbon and oxygen isotope analyses may offer clues to this puzzle because their results identify major changes in the Earth's exosystem^{4,5}. Several studies have already dealt with carbon and oxygen isotope records at the P/T transition^{2,6–11}, but we present here empirical data derived from the brachiopod shell calcite, whose isotope composition generally corresponds to that of the sea water in which the animals lived originally^{12–15}. These data allow us to assess quantitatively the magnitude of changes in the Earth's exosystem during that critical time interval.

We sampled the Upper Permian Kapp Starostin Formation¹⁶

at a number of localities in West Spitsbergen (Fig. 1). This formation consists mainly of cherts and carbonates associated with fine-grained terrigenous deposits, which accumulated in open marine environments of a broad epicontinental sea¹⁷. The conodonts found in the section at Kapp Starostin demonstrate the latest Leonardian age of the basal part of the formation and strongly suggest a Changhsingian (latest Permian) age for the upper part of the section¹⁸. In central West Spitsbergen, deposits of the Kapp Starostin Formation pass, without any evidence for discontinuity, into the overlying marly shales of the Vardebukta Formation¹⁹, which contain, in this area, Griesbachian (early Triassic) conodonts²⁰ and molluscs²¹ and, higher up, Dienerian ammonoids²². Thus the area appears to represent one of the few continuous stratigraphic records across the P/T transition. Time correlation among individual sections within the Kapp Starostin Formation (Fig. 2) has been achieved by means of event stratigraphy²³.

The samples we analysed for carbon and oxygen isotopes are brachiopod shells with fully preserved primary microstructure and without any evidence for allochthoneity. To ensure consistency of the isotope analysis, we focused on a single brachiopod taxon, the *Productacea* (although we also took some other fossil and rock samples for comparative purposes). This material is particularly suitable for such study because brachiopods fractionate only insignificantly the carbon and oxygen isotopes they take up from the sea water for shell development, and the isotope composition of their low-magnesium calcite is generally not affected by the diagenetic alterations that so often obscure the chemistry of rock matrix and skeletal calcite^{12-15,25-27}.

The carbon and oxygen isotope composition of our samples was determined by mass spectrometry on CO₂, obtained by calcite reaction with 100% H₃PO₄; the standard error of single measurements is less than 0.08‰ (ref. 28). The results are expressed here in $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ notation (relative to the PDB standard, using the NBS-19 reference sample).

The brachiopod-bearing facies occurs in West Spitsbergen throughout the entire time interval represented by the Kapp Starostin Formation, but discontinuously in each section. Therefore, we employed the event-stratigraphic time correlation to produce composite curves of changes in the carbon and oxygen isotope ratios (Fig. 3). The productacean samples yield a very clear pattern. The supplementary samples show a little more variation, but they nevertheless largely follow the productacean pattern. As clearly demonstrated by a comparison of Figs 2 and 3, the stratigraphic sequence of carbon isotope measurements taken on samples collected from each particular section is invariably consistent with the overall pattern, so that the most striking features of the pattern are in fact documented within individual sections. This comparison ensures the reliability of our composite curves as an orderly stratigraphic record.

The lower part of the carbon isotope curve from the Kapp Starostin Formation scatters around a $\delta^{13}\text{C}$ value of +4‰, which is typical of measurements taken on a variety of Permian rocks from all over the world^{5-11,14,29,33}. In the middle part of the section, $\delta^{13}\text{C}$ increases rapidly by as much as 3.5‰, a phenomenon recorded also in many other sections^{2,5-9,30,34}. The upper part of the curve shows a dramatic decline of $\delta^{13}\text{C}$ by more than 10‰, falling to pronouncedly negative values. A distinct drop in the carbon isotope ratio at the P/T transition occurs also in Transcaucasian², Chinese¹⁰ and Alpine sections¹¹, but nowhere has it been observed to be of such extraordinary magnitude. Thus, the carbon isotope curve obtained in West Spitsbergen is generally consistent with all the data reported previously. Given the biostratigraphic data we have, the time-span of the two carbon isotope shifts together is of the order of a few million years. The oxygen isotope curve essentially mimics the carbon curve, although its decline in the uppermost part of the Kapp Starostin Formation is slightly less dramatic, presumably due to a regional cooling³⁵.

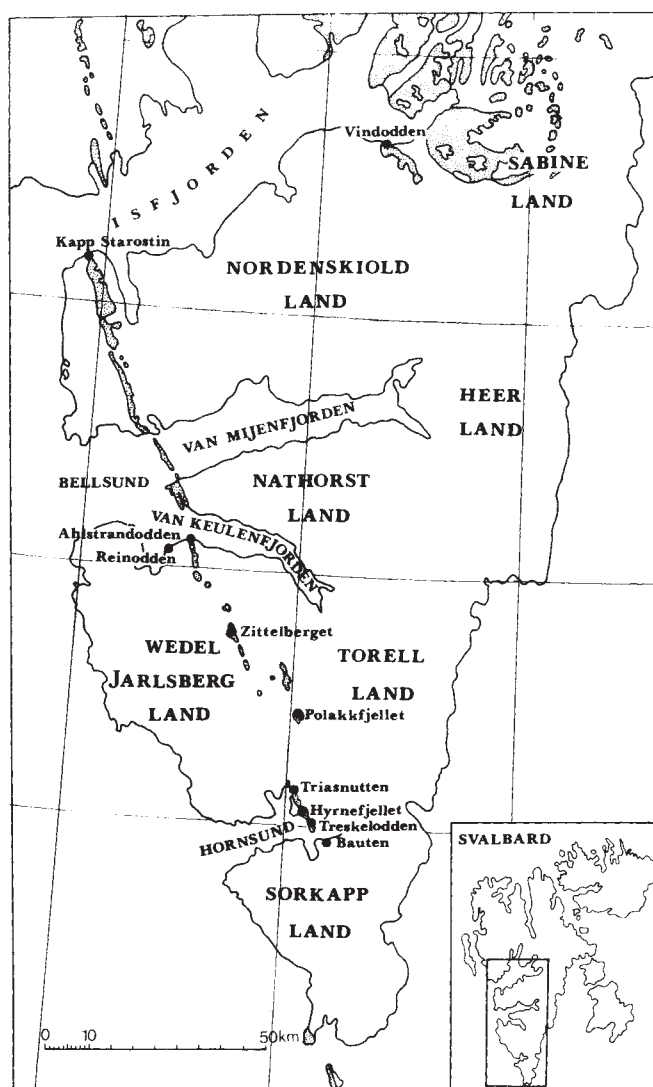


Fig. 1 Location map of the outcrops (shaded areas) of the Kapp Starostin Formation in West Spitsbergen.

To ensure that these carbon and oxygen isotope patterns are not biased by diagenetic alterations, we tested our brachiopod shell material for minor-element proportions. The results of this analysis are presented in Fig. 4 and show that our brachiopod calcite material has not been substantially altered (compare refs 14 and 15); they also show that the samples with extreme $\delta^{13}\text{C}$ values, which essentially determine the isotope patterns, do not significantly deviate from the others in this respect. It is worth noting, moreover, that even severe diagenetic changes under the influence of meteoric water have not been reported to effect a change of more than 2–4‰ in $\delta^{18}\text{O}$, as compared to the 7‰ shift we observe, and could not possibly lead to the $\delta^{18}\text{O}$ values as extreme as –11‰ that we find³⁶⁻³⁸. Having thus ruled out the possibility of diagenetic controls on our empirical pattern, we must interpret the covarying curves of $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ as reflecting a global process in the Earth's exosystem.

In this perspective, the magnitude of the carbon isotope shifts is the most extraordinary feature of the pattern. The carbon isotope curve implies that the Late Permian ocean was at first relatively impoverished (spike) and then enriched (drop) in the light isotope of carbon. Relative impoverishment in the light isotope indicates removal of isotopically light organic carbon

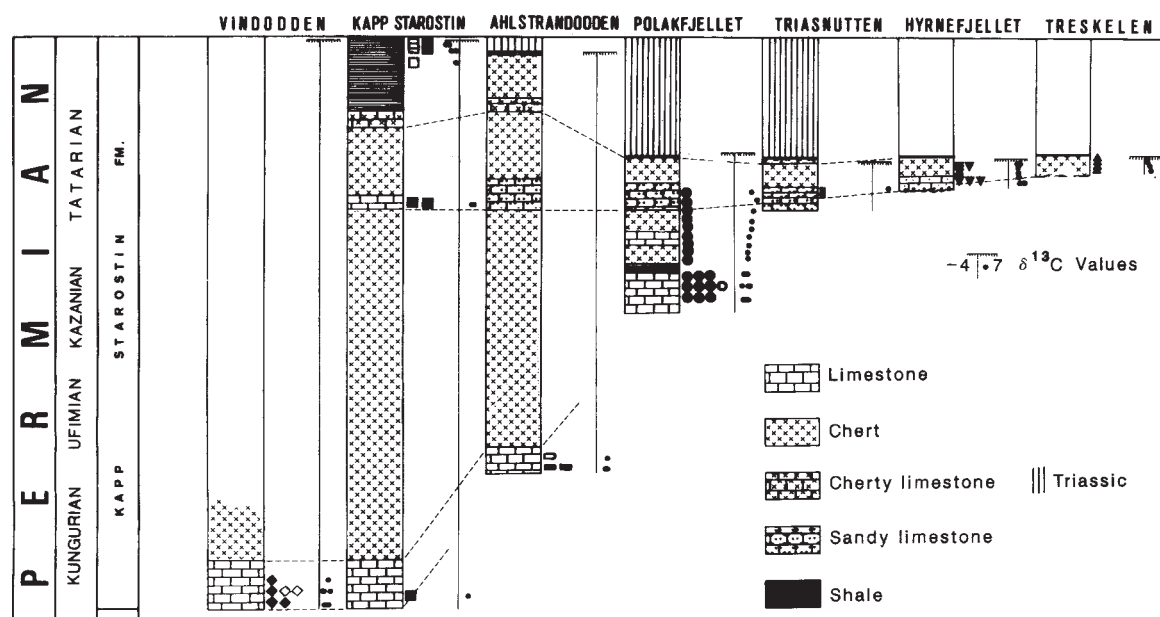
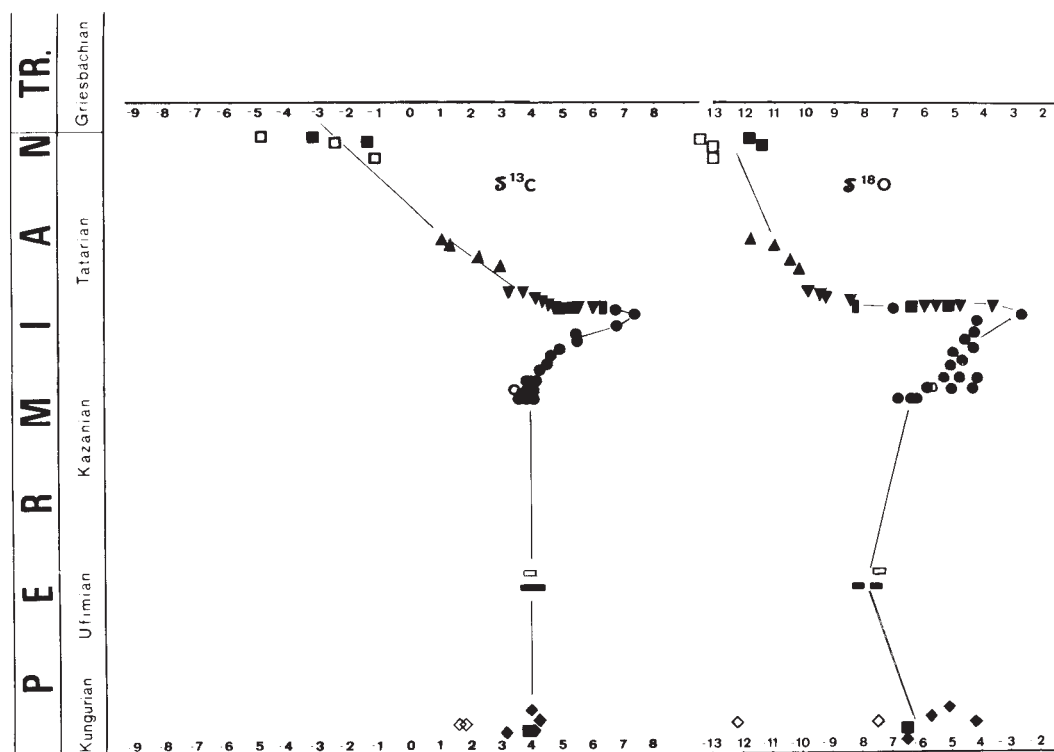


Fig. 2 Time correlation of the investigated sections and sample location; closed symbols denote productacean brachiopod shell material, open symbols indicate other fossil and bulk rock material; $\delta^{13}\text{C}$ values for individual samples are shown graphically; bold lines in the bar charts denote the P/T boundary; dashed lines are time correlation lines.

Fig. 3 Carbon and oxygen isotope curves obtained for the Kapp Starostin Formation; sample symbols are as in Fig. 2.



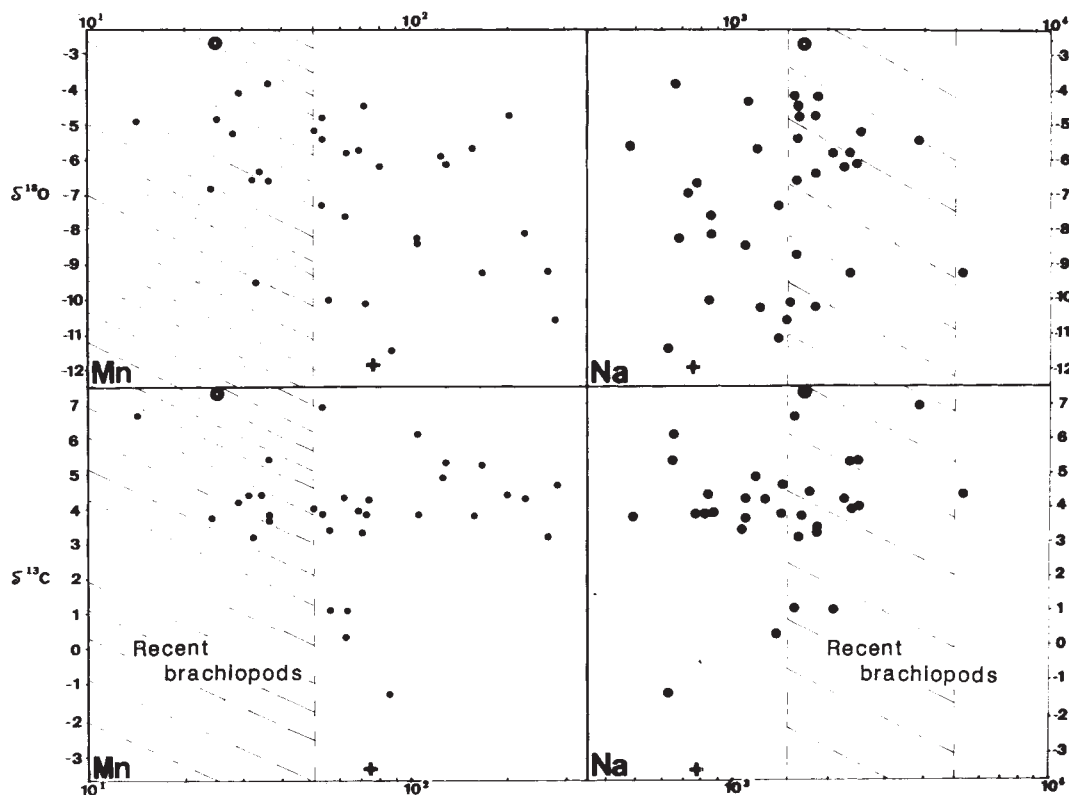
from the ocean-atmosphere system, that is, increased burial rates; relative enrichment, in turn, may be explained either by diminished burial rates of organic carbon, by diminished uptake of the light isotope by the biosphere, or by its increased input to the system, in the form of organic carbon from land ($\delta^{13}\text{C}$ of -25%), marine deposits (-15 to -20%), or as juvenile, volcanic carbon dioxide ($\delta^{13}\text{C}$ of -5%). Spitz and Degens³⁹ have derived a formula to calculate the amount of carbon with a particular isotopic composition that must be supplied to, or removed from, the reservoir with a given isotope ratio in order

to cause a specified shift in $\delta^{13}\text{C}$:

$$N_B = N_A \cdot \frac{\left[1 + \delta_{\text{std}} \cdot \left(1 + \frac{\delta_B}{1,000} \right) \right] (\delta_A - \delta_M)}{\left[1 + \delta_{\text{std}} \cdot \left(1 + \frac{\delta_A}{1,000} \right) \right] (\delta_M - \delta_B)}$$

where N_A is the initial reservoir mass of carbon, N_B is the perturbing mass and δ_s denote the isotopic compositions in standard notation, with subscripts indicating the initial (A), the

Fig. 4 Manganese and sodium proportions in the productacean shell samples analysed for carbon and oxygen isotopes, as compared with the values typically obtained in Recent brachiopods, show no diagenetic trends in the material (cf. refs 14 and 15); samples with extremely positive (circle) and extremely negative (cross) $\delta^{13}\text{C}$ are indicated. Shaded regions include Recent brachiopods.



perturbing (B), the resulting (M) and the reference standard (Std) values. From this equation and assuming that the Late Permian ocean contained as much carbon as the present one (4×10^{13} tonnes), the amount of volcanic carbon necessary to cause the latest Permian drop in $\delta^{13}\text{C}$ would demand an increase in juvenile carbon dioxide supply by three orders of magnitude, which is implausible. Hence, organic carbon must have been shifted from one reservoir to another. Assuming that the organic carbon shifted had a $\delta^{13}\text{C}$ value of -15% and that the Late Permian biosphere contained as much carbon as the present one (5.6×10^{11} tonnes), the positive spike in $\delta^{13}\text{C}$ which we observe is equivalent to removal from the ocean-atmosphere system of more than ten times the total amount of carbon in the living biosphere; the subsequent drop in $\delta^{13}\text{C}$ is equivalent to input of organic carbon in amounts comparable to the total carbon content in the ocean, or two orders of magnitude more than the total living biomass. Neither a complete cessation of carbon uptake by the biosphere (as a result of its total extinction) nor a dramatic decrease in organic matter burial could account for such a rise in the amounts of isotopically light carbon contained in the ocean.

Hence, the global process reflected by our carbon isotope curve can only be envisaged as increased burial of organic matter (spike) followed by its oxidation (drop). This interpretation is indeed consistent with our oxygen isotope curve, because oxidation of organic matter, reflected by a drop in $\delta^{13}\text{C}$, must lead—through kinetic fractionation processes—to a decline in $\delta^{18}\text{O}$ as measured in all species of dissolved oxidized carbon⁴⁰.

As we deal here with changes in isotope composition of the sea water, the most efficient explanation for our empirical pattern is that both the initial burial and the subsequent oxidation of organic matter occurred in the ocean, rather than in another carbon reservoir. A similar phenomenon took place at the Precambrian/Cambrian transition⁴¹. This phenomenon can be explained by oxidation of organic matter accumulated previously in the ocean simply because continents were barren at the time and could not yet provide large amounts of organic carbon. Our data thus indicate very high depositional rates of

organic matter in the Late Permian ocean, which were replaced at the P/T transition by its extensive oxidation.

This conclusion has tremendous implications for the history of the ocean-atmosphere system and, consequently, life. The increased burial in the ocean must have left much oxygen free. Under the resultant reducing conditions at the sea bottom, phosphorus and nitrogen recycling was enhanced⁴². The subsequent oxidation of organic matter, however, must have drawn much oxygen from the atmosphere, and oceanic phosphorus and nitrogen then sank to the sea floor because of their oxidation⁴³. The resulting decline in atmospheric oxygen levels and nutrient deficiency could be the main causes of mass extinction at the P/T transition.

Received 15 April; accepted 11 November 1988.

1. Jablonski, D. *Dynamics of Extinction* (ed. Elliott, D. K.) 183–230 (Wiley, New York, 1986).
2. Holser, W. T. & Magaritz, M. *Modern Geol.* **11**, 155–180 (1987).
3. Hoffman, A. *Arguments on Evolution* (Oxford Univ. Press, New York, 1988).
4. Hoefs, J. *Am. geophys. Un. Monogr.* **5**, 110–119 (1981).
5. Berner, R. A. *Am. J. Sci.* **287**, 177–196 (1987).
6. Magaritz, M. *et al. Earth planet. Sci. Lett.* **66**, 111–124 (1983).
7. Clemmensen, L. B., Holser, W. T. & Winter, D. *Geol. Soc. Denmark Bull.* **33**, 253–260 (1985).
8. Holser, W. T. & Magaritz, M. *Geol. Bundesanst. Wien Jb.* **128**, 75–82 (1985).
9. Holser, W. T., Magaritz, M. & Clark, D. L. *Am. J. Sci.* **286**, 390–402 (1986).
10. Xu, D. *et al. Modern Geol.* **10**, 365–376 (1986).
11. Magaritz, M. *et al. Nature* **331**, 337–339 (1988).
12. Lowenstam, H. A. *J. Geol.* **69**, 241–260 (1961).
13. Popp, B. N., Anderson, T. F. & Sandberg, P. A. *J. sedim. Petrol.* **56**, 715–726 (1986).
14. Veizer, J., Fritz, P. & Jones, B. *Geochim. cosmochim. Acta* **50**, 1679–1696 (1986).
15. Al-Aasm, I. S. & Veizer, J. *J. sedim. Petrol.* **52**, 1101–1109 (1982).
16. Cutbill, J. L. & Challinor, A. *Geol. Mag.* **102**, 418–439 (1965).
17. Małkowski, K. & Hoffman, A. *Acta Palaeont. pol.* **24**, 217–230 (1979).
18. Szaniawski, H. & Małkowski, K. *Acta Palaeont. pol.* **24**, 231–264 (1979).
19. Birkenmajer, K. *Studia Geol. pol.* **44**, 7–33 (1977).
20. Hatleberg, E. E. & Clark, D. L. *Geol. et Palaeont.* **18**, 101–126 (1984).
21. Nakazawa, K., Nakamura, K. & Kimura, G. *Proc. Japan Acad. Sci.* **63**, 171–174 (1987).
22. Gaździcki, A. & Trammer, J. *Acta Geol. pol.* **27**, 349–356 (1977).
23. Małkowski, K. *Palaeont. pol.* **43**, 69–81 (1982).
24. Veizer, J. & Fritz, P. *Earth planet. Sci. Lett.* **33**, 255–260 (1976).
25. Brand, U. & Veizer, J. *J. sedim. Petrol.* **50**, 1219–1236 (1980).
26. Brand, U. & Veizer, J. *J. sedim. Petrol.* **51**, 987–997 (1981).
27. Brand, U. *Chem. Geol.* **31**, 1–28 (1981).
28. Halas, S. & Skórzynski, Z. *J. Phys. E* **14**, 509–512 (1981).
29. Veizer, J., Holser, W. T. & Wilgus, C. K. *Geochim. cosmochim. Acta* **44**, 579–587 (1980).
30. Magaritz, M. & Schulze, K. H. *Contr. Sedi.* **9**, 269–277 (1980).
31. Rao, C. O. & Green, D. C. *J. sedim. Petrol.* **52**, 111–126 (1982).

32. Popp, B. N., Anderson, T. F. & Sandberg, P. A. *Bull. geol. Soc. Am.* **97**, 1262–1269 (1986).
33. Beauchamp, B., Oldershaw, A. E. & Krouse, K. R. *Chem. Geol.* **65**, 391–414 (1987).
34. Magaritz, M. & Turner, P. *Nature* **297**, 389–390 (1982).
35. Gruszczynski, M. & Malkowski, K. *Polish Polar Res.* **8**, 201–215 (1987).
36. Allan, J. R. & Matthews, R. K. *Sedimentology* **29**, 797–818 (1982).
37. Brand, U. *Sedimentology* **29**, 139–148 (1982).
38. Wagner, P. D. & Matthews, R. K. *J. sedim. Petrol.* **52**, 3–18 (1982).
39. Spitz, A. & Degens, E. T. *Mitt. Geol. Paläont. Inst. Univ. Hamburg* **59**, 155–166 (1985).
40. Hoefs, J. *Stable Isotope Geochemistry* (Springer, Berlin, 1987).
41. Aharon, P., Schidlowski, M. & Singh, I. B. *Nature* **327**, 699–702 (1987).
42. Berner, R. A. *Limnol. Oceanogr.* **22**, 781–786 (1977).
43. Froelich, P. N. *et al. Am. J. Sci.* **282**, 474–511.

Peat acidification in Scotland

U. Skiba*, M. S. Cresser*, R. G. Derwent†
& D. W. Futtý‡

* Department of Plant and Soil Science, University of Aberdeen,
Aberdeen AB9 2UE, UK

† Modelling and Assessments Group, Environmental and Medical
Sciences Division, Harwell Laboratory, Oxfordshire OX11 0RA, UK

‡ Land Use Division, The Macaulay Land Use Research Institute,
Craigiebuckler, Aberdeen AB9 2QJ, UK

Simulation experiments have shown that soils can become more acid when treated with accelerated loads of acid precipitation^{1–3}. Field observations, however, give less clearcut answers. Acid deposition is rarely the only variable between sites, and effects are too small for soil pH changes to be measurable over much less than a decade. To circumvent these problems, Hallbacken and Tamm⁴ in Sweden and Billett *et al.*⁵ in Scotland re-sampled soil profiles in forests after 39–57 years and showed that many surface organic and mineral horizons had acidified, but biological acidification and acid deposition may both be implicated in the observed changes⁵. Here we compare peat acidity and base saturation with modelled acid deposition at 123 sites in Scotland. Peats with highest acidity ($\text{pH}(\text{CaCl}_2) \leq 3.0$) and with lowest base saturation ($\leq 10\%$) are found mainly where deposited acidity was greater than $0.8 \text{ kg H}^+ \text{ ha}^{-1} \text{ yr}^{-1}$, providing conclusive evidence that acid deposition plays an important part in the acidification of some soils.

Dystrophic and dystrophic-flushed peats are believed to be particularly susceptible to acid deposition. The surface layers of such peats rely directly upon ion-exchange properties and precipitation for the supply of cations and anions. Acid precipitation increases the leaching of cations from organic soils³; exchangeable cations are replaced by H^+ from the incoming precipitation. For soils with only thin organic layers, leached cations may be replaced by biogeochemical weathering in the underlying mineral layers and by elemental cycling. Ions may be transported upwards from the mineral soil into the organic layer by roots penetrating into the underlying deeper horizons, and also by periodic rises of the water table. In soils such as peat, in which the organic horizons are thick ($\geq 0.5 \text{ m}$), this mechanism works only to a very limited extent⁶. Therefore the surface of dystrophic and dystrophic-flushed peats should reflect deposition chemistry rather than underlying soil mineralogy, and can be a valuable tool in estimating the effects of rain chemistry on peat chemistry.

We have compared the peat chemistry and a model of acid deposition at 123 sites in Scotland. Data on the chemistry of hill peats were extracted from the Scottish National Soil Inventory set up by the Macaulay Land Use Research Institute (Aberdeen). Soil samples for this inventory were collected after 1978. They were air-dried at 30°C , and 2-mm-sieved samples were stored until analysis, usually within a year. Soils are classified as peat when the organic matter was at least 0.5 m thick. Only dystrophic and dystrophic-flushed peats, collected from sites where the slope form was straight or convex, were included in this study. Most peat samples were collected from the top 15 cm

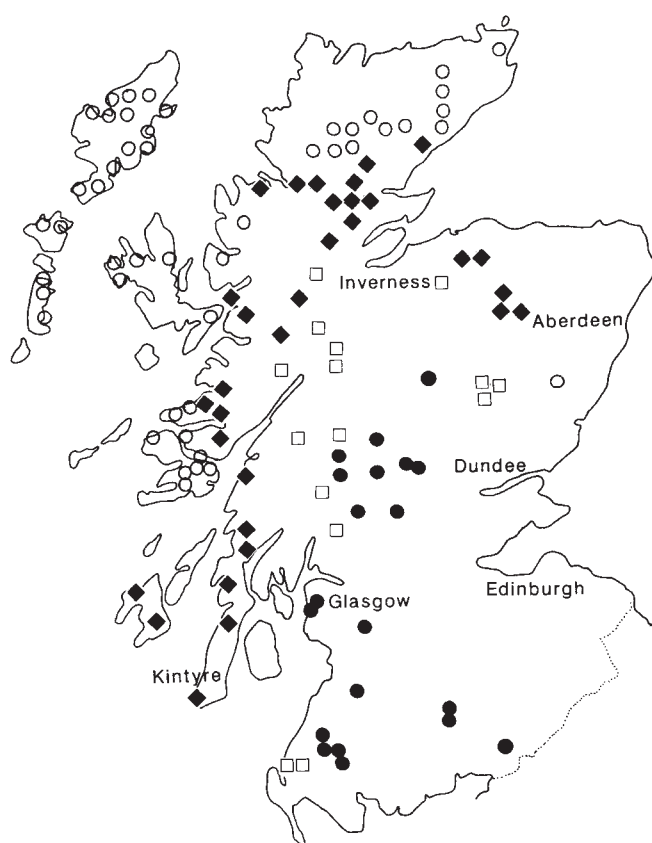


Fig. 1 Modelled wet- and dry-deposited acidity from S, N and Cl species in Scotland ($\text{kg H}^+ \text{ ha}^{-1} \text{ yr}^{-1}$). Symbols are: ○, ≤ 0.6 ; ◆, $>0.6–0.8$; □, $>0.8–1.0$; ●, >1.0 .

of the peat horizon, some from the top 20 cm and three samples from the top 25 cm. Samples were analysed for pH (CaCl_2) and percentage of base saturation at the Macaulay Institute, using standard methods. Peat pH and base saturation were plotted and their distribution in Scotland was compared with total wet- and dry-deposited acidity calculated for each peat sampling site using the Harwell trajectory model⁷. Deposited acidity at each sampling site was estimated by including contributions from both dry- and wet-deposition routes, and from acidic sulphur (SO_x), nitrogen (NO_x) and chlorine (HCl) pollutants emitted from sources within the United Kingdom and the rest of Europe. Total wet- and dry-deposited acidity was expressed in $\text{kg H}^+ \text{ ha}^{-1} \text{ yr}^{-1}$, assuming that 2 g H^+ are equivalent to 32 g S , 1 g H^+ to 14 g N and 1 g H^+ to 35.5 g Cl , and making no adjustment for the deposition of ammonia and its derivatives.

The total acid deposition rate in the trajectory model decreased by more than a factor of three across Scotland, from a maximum of $>1 \text{ kg H}^+ \text{ ha}^{-1} \text{ yr}^{-1}$ in south Scotland to a minimum of $<0.6 \text{ kg H}^+ \text{ ha}^{-1} \text{ yr}^{-1}$ in the Outer Hebrides (Fig. 1). Although the prevailing winds throughout Scotland are dominated by winds from the north-westerly through to south-westerly sectors, much of the wet- and dry-deposited acidity in the model is associated with winds and rain from the easterly through to southerly sectors. This reflects the important contribution to deposited acidity from emission sources in England and in the rest of Europe.

Most peats in the areas where acid deposition was $>0.8 \text{ kg H}^+ \text{ ha}^{-1} \text{ yr}^{-1}$ had low pH values ($\leq \text{pH } 3$, Fig. 2) and low base saturation ($\leq 10\%$, Fig. 3). The correlation coefficients (r^2) for peat pH–total acid deposition and for base saturation–total acid deposition were significant, at $p < 0.02$ ($r^2 = 0.051$) for the former and $p < 0.001$ ($r^2 = 0.278$) for the latter.

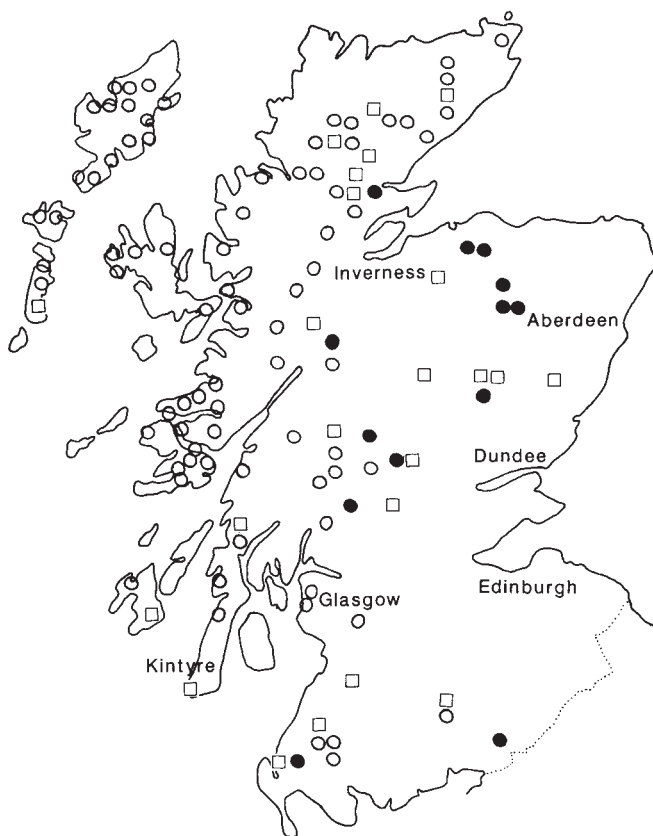


Fig. 2 The pH (CaCl_2) of peats in Scotland. ●, ≤ 2.8 (mean pH of all sites in this group analysed was 2.75); □, > 2.8 – 3.0 (mean pH 2.94); ○, > 3.0 (mean pH 3.35).

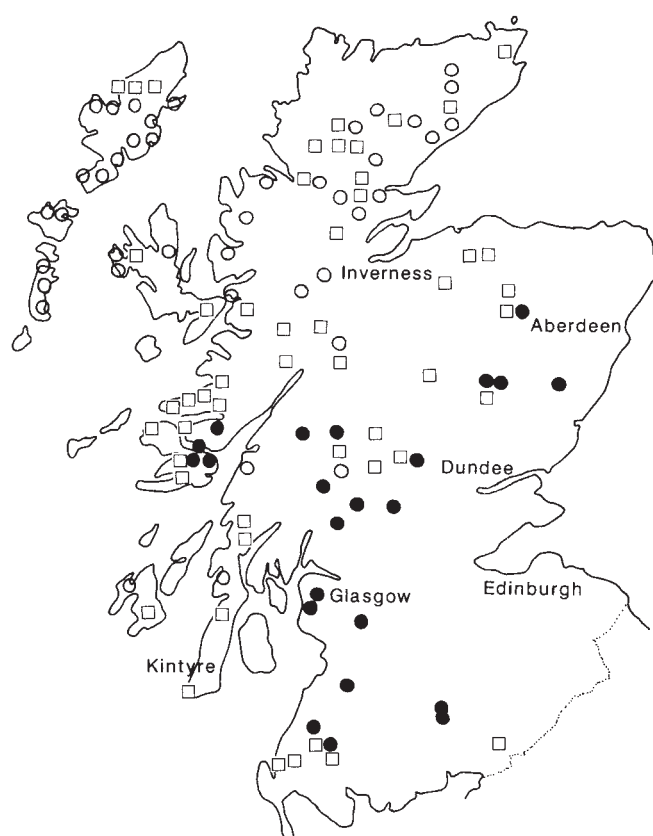


Fig. 3 Percentage base saturation in peats in Scotland. ●, $\leq 10\%$ (mean % base saturation 6.4); □, $> 10\%$ – 20% (mean % base saturation 15.6); ○, $> 20\%$ (mean % base saturation 25.1).

Peats were most acidic ($\text{pH} \leq 2.8$) between Glasgow and Inverness. Acid deposition, however, was greatest in the south, but the number of acid peats sampled here was lower because the peat is less extensive in this area. Land use and management, and position of the peats sampled on a hill site, influence the chemistry of the peat and contribute to the variability in the data observed. Some acid peats (between pH 2.8 and 3.0) were also found on Kintyre and north of Inverness. The spatial distribution of acidifying emissions shows that these pockets of high peat acidity coincided with pockets of SO_2 and NO_x emissions greater than those in surrounding areas⁸.

Acid deposition and low base-saturation distributions (Figs 1 and 3) show very good agreement. Low base saturation did not correlate with other variables, such as altitude or rainfall amount. 70% of all peats sampled with a base saturation of $< 10\%$ came from areas where the modelled wet- and dry-deposited acidity was $> 0.8 \text{ kg H}^+ \text{ ha}^{-1} \text{ yr}^{-1}$. On the other hand, 60% of all peats with a base saturation greater than 20% were collected from areas where the deposited acidity was $< 0.6 \text{ kg H}^+ \text{ ha}^{-1} \text{ yr}^{-1}$.

The distribution of acid peats and low base saturation of peats in Scotland also correlates well with the recent loss of acid-sensitive lichens. Some lichens, particularly *Lobaria pulmonaria*, are sensitive to low levels of air pollution, primarily SO_2 . Distribution maps of *Lobaria pulmonaria* showed that, since 1960, it has disappeared from sites in eastern Scotland, around Edinburgh and Glasgow and also at two sites north of Inverness⁹.

As emissions of acid-producing pollutants are declining in many European countries, it is of interest to predict how long it will take for acidified soils to become less acid. A simulation experiment on an accelerated timescale suggests that a decrease in peat acidity should be noticed in a relatively short time

period—years rather than decades³. Assuming a precipitation pH of 5, extrapolation of published results³ of simulation experiments showed that a peat now having a pH of 2.8 should decrease in acidity by 0.34 pH units within 10 years, assuming a rainfall of 2,000 mm. After 100 years a decrease in acidity of 0.53 pH units can be expected. The present difference in peat acidity between the south and east of Scotland and the north-west of Scotland is 0.55 pH units (Fig. 2). This suggests that simulated rain experiments yield realistic values for possible pH changes occurring.

The significant correlation between peat acidity or base saturation and total acid deposition gives strong evidence for acid deposition having caused further acidification of peats in Scotland over and above natural levels. As emissions of acid-producing pollutants are declining in many European countries, recovery should become evident in a few decades.

The authors are indebted to the UK Department of the Environment for financial support for this work, and to Dr M. N. Court for assistance with statistics.

Received 30 August; accepted 17 November 1988.

1. Abrahamsen, G., Bjor, K., Horntvedt, R. & Tvieta, B. in *Impact of acid precipitation on forest and freshwater ecosystems in Norway* (ed. Brackke, F. H.) 33–63 (F.R. 6/76, Oslo-As, Norway, 1976).
2. Tamm, C. O. *Sveriges SkogsFor. Tidskr.* 75, 189–200 (1977).
3. Skiba, U. & Cresser, M. S. *Wat. Res.* (in the press).
4. Hallbäck, L. & Tamm, C. O. *Scand. J. For. Res.* 1, 219–232 (1986).
5. Billett, M. F., FitzPatrick, E. A. & Cresser, M. S. *Soil Use Manage.* 4, 102–107 (1988).
6. Heal, O. W. & Smith, R. A. H. in *Production Ecology of British Moors and Montane Grasslands* (ed. Heal, O. W. & Perkins, D. F.) 3–16 (Springer, Berlin, 1978).
7. Derwent, R. G., Dollard, G. J. & Metcalfe, S. E. *Q. J. R. met. Soc.* 114, 1127–1152 (1988).
8. UK Review Group on Acid Rain, *Acid Deposition in the United Kingdom 1981–1985*, Warren Spring Laboratory, Stevenage.
9. Looney, J. H. H. & James, P. W. in *Acid Rain and Britain's Natural Ecosystems* (eds Ashmore, M., Bell, N. & Garretty, C.) 13–25 (Imperial College, London, 1988).

Female transfer and inbreeding avoidance in social mammals

T. H. Clutton-Brock

Large Animal Research Group, Department of Zoology,
Cambridge CB2 3EJ, UK

In most social mammals, males leave their natal group to breed in other groups whereas females commonly remain in the same group throughout their lives¹. In a few species however, females usually transfer between groups during adolescence¹. The functional significance of sex differences in dispersal and their connection, if any, to the avoidance of inbreeding is disputed²⁻⁷. Here I show that in polygynous mammals where females commonly remain to breed in their natal group, their average age at first conception typically exceeds the average period of residence of adult males in breeding groups. In contrast, where females usually transfer to breed in other groups, the average residence of breeding males or of resident male kin groups typically exceeds the average age of females at first conception. These results support the suggestion that female mammals commonly transfer to avoid inbreeding with their father or other close relatives, although female dispersal may also occur for other reasons^{8,9}.

Most polygynous, group-living mammals can be classified into three groups on the basis of dispersal. In the majority of species, most males emigrate from their natal group at or shortly after puberty and transfer to other breeding groups, either immediately or at some later stage of their lifespan¹. Although female emigration is not uncommon in some of these species and female groups may subdivide and occasionally coalesce³, female transfer into established breeding groups is unusual. Second, in a small number of species females habitually transfer to other breeding groups during adolescence, whereas males commonly remain in their natal group¹⁰⁻¹⁵. Finally, in a few

species that live in uni-male harem groups, both sexes usually disperse from their natal group at adolescence to breed in other groups¹⁶⁻¹⁹.

Male-biased dispersal may predominate in polygynous mammals because adolescent males are commonly ejected from their natal group by resident males^{2,3}, because the potential benefits of dispersal are usually higher for males than females²⁰⁻²² and because the costs of dispersal are generally higher for females because of their lower age at first breeding⁷, their more stringent energetic requirements^{5,23}, and their dependence on coalitions with relatives²⁴. Where the majority of males disperse at adolescence, females will seldom need to disperse to avoid inbreeding. So why do females usually disperse in some species? One possible explanation is that females habitually transfer during adolescence in species where the effective breeding lifespans of males within particular groups commonly exceed the age of females at first conception^{1,4,21} and where females would risk inbreeding with their father or other close relatives if they remained in their natal group. If so, predominant female dispersal should be associated with male residence in breeding groups that usually exceeds the age of females at first breeding, whereas in species where females remain in their natal group, male residence should usually be shorter than the age at which females breed for the first time.

To test this prediction, I extracted all available estimates of the mean age of females at first conception and of the mean residence of breeding males in particular groups or breeding territories for social, polygynous mammals from the literature. Where several related males defend access to female groups, I used the minimum residence of male kin groups. For one species (grey kangaroos) where a single male is responsible for most matings, I used the average period for which males held alpha rank. The eventual sample included eighteen species where males habitually transfer from their natal group and females usually remain there (Table 1) and nine species where females commonly transfer from their natal group to breed elsewhere (Table 2). The sample of male-dispersing species included ten usually found in uni-male harems (which may include more

Table 1 Social mammals where male transfer predominates

Species	Predominant transferring sex	Mean female age at first conception (months)	Mean male residence (months)	Does female age at first breeding exceed average male residence in a single group?	Refs
Uni-male groups or territories					
Yellowbellied marmot (<i>Marmota flaviventris</i>)	M	36	27 R	Yes	29
Blacktailed prairie dog (<i>Cynomys ludovicianus</i>)	M	24	<24 R	Yes	27
Red deer (<i>Cervus elaphus</i>)	M	40	36 R	Yes	30
Wedge-capped capuchin (<i>Cebus olivaceus</i>)	M	72	120 R	No	51
Redtail monkey (<i>Cercopithecus ascanius</i>)	M	24	<24 R	Yes	31
Campbell's guenon (<i>C. campbelli</i>)	M	36	<34 R	Yes	31
Blue monkey (<i>C. mitis</i>)	M	66	<40 R	Yes	31
Gelada baboon (<i>Theropithecus gelada</i>)	M	54	43 R	Yes	32
Grey langur (<i>Presbytis entellus</i>)	M	36	28 R	Yes	33
Purple leaf monkey (<i>P. senex</i>)	M	48	36 R	Yes	34
Multi-male groups					
Eastern grey kangaroo (<i>Macropus giganteus</i>)	M	24	12 α	Yes	35
African lion (<i>Panthera leo</i>)	M	38	26 R	Yes	24
Ringtail lemur (<i>Lemur catta</i>)	M	30	<24 R	Probably	23, 36
Vervet monkey (<i>Cercopithecus aethiops</i>)	M	33	32 R	Yes	37
Yellow baboon (<i>Papio cynocephalus</i>)	M	72	<51 R	Yes	38
Rhesus macaque (<i>Macaca mulatta</i>)	M	66	19 R	Yes	39
Japanese macaque (<i>M. fuscata</i>)	M	54	<30 R	Yes	40, 41
Toque macaque (<i>M. sinica</i>)	M	60	50 R	Yes	37, 42, 43

For mean male residence: R, residence; α , tenure of α rank. Field data were used to estimate the age of females at first conception wherever possible supplemented by data from ref. 44. Average male residence was the estimated duration of reproductive activity in males within a particular group. In the majority of species, this was equivalent to mean residence per group in non-natal breeding groups but for grey kangaroos I have used the mean tenure of alpha rank and for red deer the average breeding lifespan.

Table 2 Social mammals where female transfer predominates or where both sexes usually transfer

Species	Predominant transferring sex	Mean female age at first breeding (months)	Mean male residence (months)	Does female age at first breeding exceed average male residence in a single group?	Refs
Uni-male groups					
Plains zebra (<i>Equus burchelli</i>)	M + F	24	>60 R	No	16, 45
Feral horse (<i>E. caballus</i>)	M + F	25	38 R	No	46
Hamadryas baboon (<i>Papio hamadryas</i>)	M + F	52	>36 R	No	18*
Mountain gorilla (<i>Gorilla gorilla beringei</i>)	M + F	119	180 R	No	19†
Multi-male groups					
Cape hunting dog (<i>Lycaon pictus</i>)	M + F	38	Resident male kin group	No	10, 11, 47
Red howler monkey (<i>Alouatta seniculus</i>)	M + F	49	71 R	No	8, 48‡
Red colobus monkey (<i>Colobus badius</i>)	F	46	Resident male kin group	No	12, 13, 49
Chimpanzee (<i>Pan troglodytes</i>)	F	120+	Resident male kin group	No	14, 50
Bonobo (<i>P. paniscus</i>)	F	120+	Resident male kin group	Probably not	14, 50

In four species, groups of related males defend access to groups or communities of females. The average residence of male kin groups is unknown in all cases, but take-overs are rare and the average tenure of male groups is evidently long. In one species (hamadryas baboons) females leave their natal units but remain within a larger social unit, consisting of several harems, often belonging to related males.

* Also, H. Kummer, personal communication; † A. H. Harcourt, personal communication; ‡ R. Rudran (unpublished).

than one adult male but where one is responsible for most of all mating) and eight usually found in multi-male, multi-female groups (where more than one mate is involved in breeding). The sample of female dispersing species included four species normally found in uni-male harem groups and four species where a resident kin group of males defends breeding access to a female group. Two further species, black spider monkeys *Ateles paniscus* and woolly monkeys *Lagothrix lagotricha*, may also have breeding systems of this kind but their pattern of dispersal is not yet known^{25,26}.

In all but one of the eighteen species where females typically remain in their natal group and males usually transfer to breed in other groups, the average age of females at first breeding exceeds the average residence of males in breeding groups (see Table 1). In contrast, in all nine species where females habitually transfer from their natal group, the average residence of breeding males is longer than the age at which females begin to breed (see Table 2). The residence of male kin groups is not known for hunting dogs, red colobus, chimpanzees and bonobos, but take-overs by extra-group males are evidently unusual and it is clear that the average residence of kin groups typically exceeds the age of females at first conception. There is an interesting contrast between these species and lions (see Table 1), where groups of related males also defend breeding access to female prides but take-overs are frequent, male residences are short (average 26 months) and females rarely disperse²⁴.

The distribution of dispersal thus suggests that where males can effectively defend mating access to feeding territories or to female groups for periods longer than the age of females at first breeding, females may disperse to avoid mating with close relatives. Where male dispersal at adolescence is associated with long male residence (as in gorillas) both sexes are likely to disperse from natal groups. The association between female dispersal and long male tenure does not indicate that females never disperse for other reasons: in at least two social primates, females commonly disperse in circumstances where close inbreeding is unlikely^{8,9}. Nor does it indicate that females never risk inbreeding with their fathers or brothers. In some species where females remain in their natal group, the difference between the average duration of male residence and the average age of females at first breeding is small (see Table 1) and a proportion of females will reach breeding age while their father is still active in their group³¹. Under these circumstances, recent evidence suggests that breeding males sometimes disperse or avoid mating with their daughters^{3,24,27} or females may dis-

perse¹⁹, attempt to mate with extra-group males²⁴ or delay conception²⁸.

I am grateful to Ann Pusey, Jim Moore, Sandy Harcourt, Paul Harvey, Dafila Scott, Phyllis Lee and Mariko Hiraiwa-Hasegawa for comments on the manuscript and to Rudi Rudran and Dan Rubenstein for access to unpublished data.

Received 17 October; accepted 24 November 1988.

- Greenwood, P. J. *Anim. Behav.* **28**, 1140–1162 (1980).
- Moore, J. & Ali, R. *Anim. Behav.* **32**, 94–112 (1981).
- Moore, J. *Int. J. Primatol.* **5**, 537–589 (1984).
- Pusey, A. E. & Packer, C. R. in *Primate Societies* (eds Smuts, B. B. et al.) 250–266 (University of Chicago Press, 1987).
- Pusey, A. E. *Trends Evol. Ecol.* **2**, 295–299 (1987).
- Wrangham, R. W. *Behaviour* **75**, 262–300 (1980).
- Johnson, C. W. *Oecologia* **69**, 626–627 (1986).
- Crockett, C. M. *Female Primates: Studies by Women Primatologists* (ed. Small, M. F.) 159–173 (Liss, New York, 1984).
- Marsh, C. W. *Nature* **281**, 568–569 (1979).
- Frame, L. H., Malcolm, J. R. & van Lawick, H. Z. *Tierpsychol.* **50**, 225–245 (1980).
- Malcolm, J. R. & Marten, K. *Behav. Ecol. Sociobiol.* **10**, 1–13 (1982).
- Strusaker, T. T. *The Red Colobus Monkey* (University of Chicago Press, 1980).
- Strusaker, T. T. & Leland, L. in *Advances in the Study of Behavior* (eds Rosenblatt, J. et al.) 159–225 (Academic, New York, 1979).
- Nishida, T. T. & Hiraiwa-Hasegawa, M. in *Primate Societies* (eds Smuts, B. B. et al.) 165–178 (University of Chicago Press, 1987).
- Wrangham, R. W. in *Primate Societies* (eds Smuts, B. B. et al.) 282–296 (University of Chicago Press, 1987).
- Ohsawa, H. *African Studies Monographs* **2**, 53–71 (1982).
- Kummer, H. *Social Organization of Hamadryas Baboons* (University of Chicago Press, 1968).
- Sigg, A., Abegglen, J. J. & Dasser, V. *Primates* **23**, 473–487 (1982).
- Stewart, K. & Harcourt, A. H. in *Primate Societies* (eds Smuts, B. B. et al.) 155–164 (University of Chicago Press, 1987).
- Clutton-Brock, T. H. & Harvey, P. H. in *Growing Points in Ethology* (eds Bateson, P. P. G. & Hinde, R. A.) 195–237 (Cambridge University Press, 1976).
- Packer, C. R. *Anim. Behav.* **27**, 1–36 (1979).
- Waser, P. M., Austad, S. N. & Keane, B. *Am. Nat.* **128**, 529–537 (1986).
- Budnitz, N. & Dainis, K. *Lemur Behavior* (eds Tattershall, I. & Sussman, R. W.) (Plenum, New York, 1975).
- Pusey, A. E. & Packer, C. R. *Behaviour* **101**, 275–310 (1987).
- Robinson, J. G. & Janson, C. H. in *Primate Societies* (eds Smuts, B. B. et al.) 69–82 (University of Chicago Press, 1987).
- Roosmalen, van M. G. M. Thesis, Rijksinstituut voor Natuurbeheer Arnhem, The Netherlands (1980).
- Hoogland, J. L. *Science* **215**, 1639–1641 (1982).
- Blouin, S. F. & Blouin, M. *Trends Evol. Ecol.* **3**, 230–232 (1988).
- Armitage, K. in *Ecological Aspects of Social Evolution* (eds Rubenstein, D. I. & Wrangham, R. W.) 303–331 (Princeton University Press, 1986).
- Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. in *Reproductive Success* (ed. Clutton-Brock, T. H.) 325–343 (University of Chicago Press, 1988).
- Cords, M. in *Primate Societies* (eds Smuts, B. B. et al.) 98–111 (University of Chicago Press, 1987).
- Dunbar, R. M. *Reproductive Decisions* (Croom Helm, London, 1988).
- Hardy, S. B. *The Langurs of Abu* (Harvard University Press, 1977).
- Rudran, R. *Folia Primatol.* **19**, 166–192 (1975).
- Jarman, P. & Southwell, C. D. in *Ecological Aspects of Social Evolution* (eds Rubenstein, D. I. & Wrangham, R. W.) 399–428 (Princeton University Press, 1986).
- Richard, A. F. in *Primate Societies* (eds Smuts, B. B. et al.) 25–33 (University of Chicago Press, 1987).
- Henzi, S. D. & Lucas, J. W. *Folia Primatol.* **33**, 220–235 (1980).

38. Altmann, J., Hausfater, G. & Altmann, S. A. in *Reproductive Success* (ed. Clutton-Brock, T. H.) 403–418 (University of Chicago Press, 1988).
39. Melnick, D. J., Pearl, M. C. & Richard, A. F. *Am. J. Primatol.* **7**, 229–243 (1984).
40. Sugiyama, Y. in *Advances in the Study of Behavior* Vol. 7 (eds Rosenblatt, J. S. et al.) 255–284 (Academic, New York, 1976).
41. Itani, J. in *Primate Socialization* (ed. Poirier, F. E.) 165–171 (Random House, New York, 1972).
42. Dittus, W. in *Socioecology and Psychology of Primates* (ed. Tuttle, R. H.) 125–151 (Mouton, The Hague, 1975).
43. Dittus, W. *Behaviour* **63**, 281–322 (1977).
44. Harvey, P. H. & Clutton-Brock, T. H. *Evolution* **39**, 559–581 (1985).
45. Klingel, H. J. *Reprod. Fert. Suppl.* **23**, 7–11 (1974).
46. Berger, J. *Wild Horses of the Great Basin* (University of Chicago Press, 1986).
47. Gittleman, J. *Am. Nat.* **127**, 744–771 (1986).
48. Crockett, C. M. & Eisenberg, J. F. in *Primate Societies* (eds Smuts, B. B. et al.) 54–68 (University of Chicago Press, 1987).
49. Struhsaker, T. T. & Leland, C. in *Primate Societies* (eds Smuts, B. B. et al.) 83–97 (University of Chicago Press, 1987).
50. Wrangham, R. W. in *Ecological Aspects of Social Evolution* (eds Rubenstein, D. & Wrangham, R. W.) 352–378 (Princeton University Press, 1986).
51. Robinson, J. G. *Behav. Ecol. Sociobiol.* **23**, 187–197 (1988).

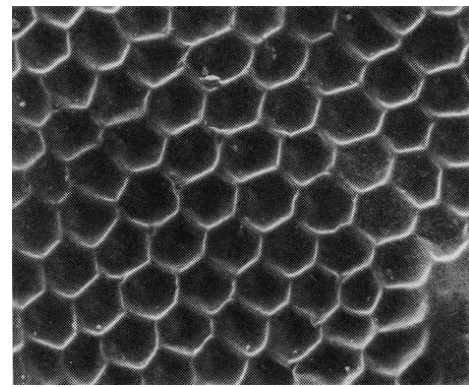


Fig. 1 Scanning electron micrograph of the elytral surface of *Cicindela oregona*. The microsculpture consists of small hexagonal alveoli that measure 13 μm across (magnification $\times 600$).

Pointillistic mixing of interference colours in cryptic tiger beetles

Thomas D. Schultz & Gary D. Bernard*

Department of Biology, Yale University, New Haven, Connecticut 06511, USA

* Department of Ophthalmology and Visual Science, Yale University School of Medicine, P.O. Box 3333, New Haven, Connecticut 06510, USA

In insects and other animals, interference colours usually serve as conspicuous visual signals^{1–4}. The colours of tiger beetles also originate by interference of multiple reflections^{5,6}, yet many species exhibit dorsal colorations that appear pigmentary and inconspicuous (cryptic) against soil backgrounds. These dull colours, however, are actually additive mixtures of different interference colours that blend when viewed with the naked eye. This same principle of partitive mixing is used in colour television, printing and the pointillist painting style of neo-impressionist artists⁷. We report here on the first demonstration of interference colours mixing partitively to produce animal camouflage.

Ultrastructural studies show that tiger beetle colours originate from a series of thin cuticular layers near the outer surface of the integument⁵. These layers function optically as a multi-layered interference filter. The layers lie directly above a thicker melanized region and conform to a surface microsculpture consisting of alveoli (hexagonal pits) that measure 13 μm across (Fig. 1). The reflectance spectrum depends on the optical thicknesses of the layers and the angles of illumination and observation⁸.

We examined the elytra (wing covers) from two colour morphs of the tiger beetle *Cicindela oregona* to determine how interference reflections produced their cryptic colorations. As in many cicindelids^{9,10}, the elytral coloration of *C. oregona* varies geographically and matches the local soil colour¹¹. The morphs exhibited brown or black elytra that appeared pigmentary to the unaided eye. When viewed through an incident-light microscope, however, the alveolar basins and ridges produced a mosaic of coloured reflections that changed with viewing angle. Moreover, discrete patches of alveoli (40–80 μm across) reflected interference colours different from those of the surrounding alveoli.

To compare colours at the macro- and microscopic levels, we measured reflectance spectra for small fields of uniform colour and larger fields that contained the different colour patches. Reflectance spectra from each colour morph were created by mounting specimens of *C. oregona* on the universal stage of an incident-light microspectrophotometer, focusing on an area representative of the elytral colour, and measuring the reflectance with monochromatic flashes at 10 nm intervals¹².

The brown morph exhibited circular patches of blue-green

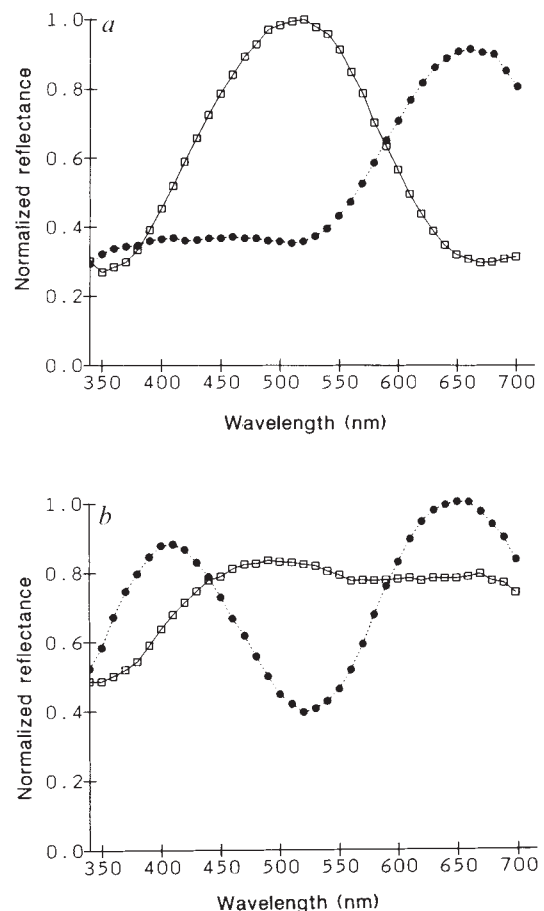


Fig. 2 Normalized reflectance spectra for adjacent microscopic colour patches on the elytra of *Cicindela oregona* measured under high power through a $\times 40/0.75$ objective. Reflectances were measured for patches 30–40 μm in diameter containing 5–8 alveoli of the same colour. The brown morph (a) contained blue-green patches (—□—) embedded in a red field (···●···). The black morph (b) exhibited irregular magenta patches (···●···) surrounded by dull green areas (—□—). The microspectrophotometer was based on a Leitz Ortholux microscope with MPV Pol-Opak illuminator, $\times 8/0.18\text{P}$ or Nikon CF Fluor $\times 20/0.75$ objectives, and Zeiss UV-projective coupled to a Hammamatsu R928 photomultiplier. The aperture stop of the Epi-illuminator was reduced to near minimum, and the incident angle was made as oblique as possible. The instrumental baseline was calibrated with a National Bureau of Standards SRM-2003D spectral reflectance standard.

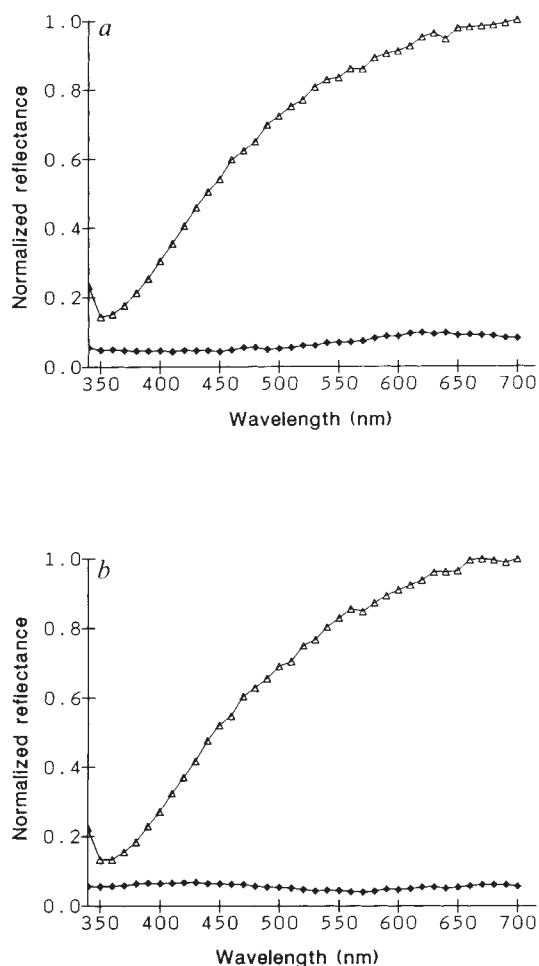


Fig. 3 Normalized reflectance spectra of coloured and white elytral areas from the same two morphs of *C. oregona* but measured under low power through a $\times 8/0.18$ objective. Reflectances were measured, for fields (300 μm in diameter) containing ~ 530 alveoli, in both the coloured and white areas of the brown (a) and black (b) elytra. Fields in the coloured area contained 7–8 microscopic colour patches described in Fig. 2. Reflectances of the coloured areas ($\triangle$) are plotted with the same normalization constant as the white areas ($\diamond$) to illustrate the desaturation and low brightness of the mixed colours.

alveoli surrounded by a red field (Fig. 2a). In the black morph, irregular patches of magenta interference colours were surrounded by dull green alveoli that reflected both green and magenta colours (Fig. 2b). The alveoli of pure magenta patches produced two prominent reflectance peaks at 660 nm and 400 nm.

Spectra produced by partitive mixing of the microscopic colour patches were measured by decreasing the magnification to include eight colour patches and their surrounding field. The resulting spectra were low in intensity and saturation when compared to an adjacent white elytral spot. The brown morph showed a broad, and slightly elevated reflectance between 500 and 750 nm due to the summation of green and red reflections (Fig. 3a). In the black elytra, the summation of green and magenta areas produced a relatively flat, unsaturated reflectance spectrum (Fig. 3b).

In tiger beetles such as *C. oregona*, reflecting layers may vary in periodicity, resulting in small clusters of alveoli that reflect different wavelengths⁵. These colour patches are too small to be resolved by the unaided eye and spatially fuse to produce an additive colour mixture¹³. Partitive mixing of different spectral combinations produces different cryptic colour morphs of *C. oregona*.

In addition, the surface microsculpture affects the intensity and wavelength content of these mixed colours. The pitted elytral surface scatters the reflected light over a wide range of angles, thereby reducing the amount of light measured by the modest collecting angle of the microscope. Similarly, much of the scattered light escapes the even smaller collecting angle of the observer's eye, causing the colours to appear dark. The reflectance spectrum is affected also by the angle of the reflecting surface relative to incident light⁸. Shorter wavelengths reflected from the angled walls of the alveoli additively mix with those reflected from the ridges and basins, reducing the purity of the colour.

Many cryptic colorations of animals are produced through partitive mixing, but most involve pigments^{14,15}. However, in tiger beetles and other Coleoptera^{16,17}, partitive mixing of different interference colours and scattering by the surface microsculpture produce broad-band reflectance spectra that match background colours and serve as camouflage.

We thank J. C. Schultz, N. F. Hadley and R. Wehner for reviewing drafts of this manuscript.

Received 26 August; accepted 7 November 1988.

1. Bleiweiss, R. *Auk* **102**, 701–713 (1985).
2. Ghiradella, H., Aneshansely, D., Eisner, T., Silberglied, R. E. & Hinton, H. E. *Science* **178**, 1214–1217 (1972).
3. Young, A. M. *Oecologia* **7**, 209–222 (1971).
4. Fogden, S. *Animals and Their Colors* (Crown, New York, 1974).
5. Schultz, T. D. & Rankin, M. A. *J. exp. Biol.* **117**, 87–110 (1985).
6. Schultz, T. D. & Hadley, N. F. *Physiol. Zool.* **60**, 737–745 (1987).
7. Williamson, S. J. & Cummins, H. Z. *Light and Color in Nature and Art* (Wiley, New York, 1983).
8. Land, M. F. *Prog. Biophys. molec. Biol.* **24**, 75–106 (1972).
9. Willis, H. L. *Univ. Kansas Sci. Bull.* **47**, 145–313 (1967).
10. Hadley, N. F., Schultz, T. D. & Savill, A. *N.Z. J. Sci.* (in the press).
11. Schultz, T. D. *Bull. ent. Soc. A* **32**, 142–146 (1986).
12. Bernard, G. D. in *Methods in Enzymology* **81** (ed. Packer, L.) 752–759 (Academic, New York, 1982).
13. Hurvich, L. M. *Color Vision* (Sinauer, Sunderland, 1981).
14. Hanlon, R. T. & Messenger, J. B. *Phil. Trans. R. Soc. B* **320**, 437–487 (1988).
15. Nijhout, H. F. *Scient. Am.* **245**, 140–151 (1981).
16. Dreisig, H. *Nature's Verden* **11**, 337–346 (1978).
17. Goulet, H. *Quaest. Ent.* **19**, 219–482 (1983).

Substitution at residue 227 of H-2 class I molecules abrogates recognition by CD8-dependent, but not CD8-independent, cytotoxic T lymphocytes

Terry A. Potter*, T. V. Rajan†, Robert F. Dick II‡ & Jeffrey A. Bluestone‡

* Division of Immunology, Department of Medicine, National Jewish Center for Immunology and Respiratory Medicine, Denver, Colorado 80206, USA

† Department of Pathology, University of Connecticut Health Center, Farmington, Connecticut 06032, USA

‡ Ben May Institute, University of Chicago, Chicago, Illinois 60637, USA

The CD8 (Lyt 2) molecule is a phenotypic marker for T lymphocytes that recognize and react with major histocompatibility complex (MHC) class I molecules. Antibody blocking experiments^{1,2} and gene transfection studies^{3,4} indicate that CD8 binds to a determinant on MHC class I molecules on the target cells, facilitating interaction between effector T lymphocytes and the target cell. The CD8 molecule may also be involved in transmembrane signalling during T-cell activation^{5,6}. The existence of CD8[−] cytotoxic T lymphocytes (CTL)⁷ and class I-reactive CTL that are not inhibited by antibody to CD8 suggests that at least some CTL do not require the CD8 molecule to interact with and lyse target cells. We have recently demonstrated that cells transfected with

an $H-2D^d$ gene that carries a mutation at residue 227 are not killed by primary CTL⁸. Here we show that although this mutation abrogates recognition by primary CTL, it does not affect recognition by CD8-independent CTL, suggesting that residue 227 of class I molecules might contribute to a determinant that is the ligand of the CD8 molecule.

We first described a somatic cell mutant isolated by immunoselection that expressed a variant $H-2D^d$ molecule which was not killed by alloreactive anti- $H-2D^d$ CTL⁹. This cell line had been isolated because of its failure to express a serological determinant expressed in the $\alpha 3$ domain. Recently⁸ we identified the mutation in the $\alpha 3$ domain of the $H-2D^d$ molecule as a Glu $\rightarrow$ Lys substitution at residue 227. This change was reproduced in a cloned $H-2D^d$ gene by oligonucleotide-directed mutagenesis. Cells transfected with this mutant gene expressed all the serological epitopes on the $H-2D^d$ $\alpha 1/\alpha 2$ domains but were not killed by anti- $H-2D^d$ CTL populations generated *in vitro* in primary responses. This was not apparently a result of the destruction of the conformation of the $H-2D^d$ molecule, because the mutant stimulated an anti- $H-2D^d$ reactive IL-2-secreting T-cell hybridoma as efficiently as the wild-type $H-2D^d$ molecule.

We first determined whether this mutation in other class I genes could abrogate CTL recognition by the creation of hybrid genes through exon shuffling^{10,11}. Hybrid genes comprised the 5' region and exons 1-3 ($\alpha 1/\alpha 2$ domains) from the $H-2K^b$ gene ligated to exons 4-8 ($\alpha 3$ domain, transmembrane and cytoplasmic regions) of the $H-2D^d$ gene (Fig. 1). One gene, referred to as $H-2K^b/D^d$ Glu, had glutamic acid at position 227 (as does wild-type $H-2D^d$), whereas the other gene, $H-2K^b/D^d$ Lys contained a mutation that leads to substitution with lysine at position 227. Both of these genes were transfected into the cell line R8.161, an $H-2^b$ variant isolated from the R8 ($H-2^b/H-2^d$) cell line. Cells expressing these genes were isolated and serological expression of $H-2K^b$ $\alpha 1/\alpha 2$ epitopes, defined by a panel of antibodies to $H-2K^b$, was confirmed by flow-microfluorimetry analysis (Fig. 2). The reactivity of antibodies to the $\alpha 1/\alpha 2$ domains of $H-2K^b$ with the transfectants expressing either the $H-2K^b/D^d$ Glu or $H-2K^b/D^d$ Lys genes was identical, indicating that this substitution in the $\alpha 3$ domain did not significantly affect the conformation of the $\alpha 1/\alpha 2$ region of $H-2K^b$. These cells were also used as targets for anti- $H-2K^b$

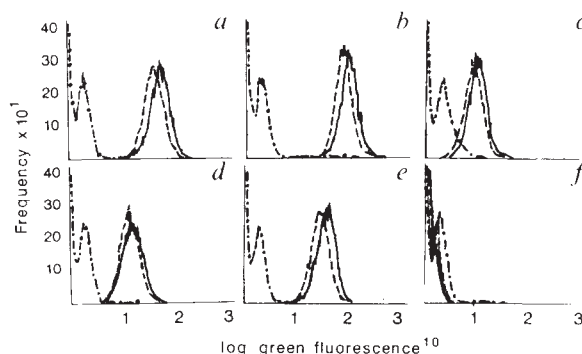


Fig. 2 Flow-microfluorimetry analysis of $H-2K^b$ expression on R8.161 and transfected cell lines.

Methods. Cells were analysed for serological expression of $H-2K^b$ epitopes by incubating 10^6 cells with $100 \mu\text{l}$ $H-2K^b$ antibodies at an optimal dilution on ice for 30 min. The cells were washed and then incubated for a further 30 min with fluorescein isothiocyanate-conjugated goat anti-mouse immunoglobulin. After staining, the cells were washed and the fluorescence of 10,000 viable cells measured on an EPICS 752B. The fluorescence profiles for R8.161 $H-2K^b$ Glu (---), R8.161 $H-2K^b$ Lys (—) and R8.161 (···) are shown. The determinants recognized by the EH144 (a) and B8-24-3 (b) antibodies have been mapped to the $\alpha 1$ domain, antibodies 5F1 (c) and 28-13-3 (d) to the $\alpha 2$ domain; residues from both $\alpha 1$ and $\alpha 2$ contribute to the determinant recognized by the K9-178 (e) antibody^{11,16}. Antibody 128-3-3 (f) is specific for the $H-2K^b$ bm^{10} mutation and is used here as a non-reactive control.

CTL generated in a primary *in vitro* response (B10.D2) (ddd) anti-B10 (bbb). These CTL lysed cells expressing the $H-2K^b/D^d$ Glu gene product (Fig. 3a), but did not lyse transfectants expressing the product of the $H-2K^b/D^d$ Lys gene (Fig. 3d). This reactivity was similar to that of anti- $H-2D^d$ CTL with cells expressing the mutant or wild-type $H-2D^d$ gene⁸. Adding antibody to CD8 inhibited the reactivity of the CTL for cells expressing the $H-2K^b/D^d$ Glu gene product (Fig. 3a). To find out the role of CD8 in the lysis of these cells, we generated CTL that were less readily inhibited by anti-CD8 antibody by priming B10.D2 mice with EL4 cells, followed one week later by restimu-

Fig. 1 Construction of hybrid genes used in this study. The $H-2D^d$ gene used here actually contains a 700-base pair deletion from the *Kpn*I to *Hpa*I site in the third intervening sequence of the $H-2D^d$ gene isolated by Margulies *et al.*¹⁴. This deletion, which was produced during the religation of the 2.1-kilobase (kb) *Hpa*I fragment containing exons 4-8 which had been subcloned into M13mp19 for site-directed mutagenesis, did not affect expression of the gene⁸. The mutation at residue 227 was produced by oligonucleotide-directed mutagenesis as previously described⁸. The $H-2K^b/D^d$ Glu gene was produced by replacement of the 1.9-kb *Xba*I fragment encoding exons 1-3 of the $H-2D^d$ gene with a similar fragment from the $H-2K^b$ gene. Digestion with *Pvu*II ensured that exons 1-3 were inserted in the correct orientation. The hybrid genes were cloned into the pSV2Neo plasmid. Plasmid DNA containing these genes was introduced into R8.161 by electroporation¹⁵. After 48 hours incubation at 37 °C, G418 antibiotic (Gibco) was added to a concentration of 1.0 mg ml^{-1} . X denotes the location of the *Xba*I sites that flank exons 1-3 of both $H-2D^d$ and $H-2K^b$ and * denotes the location of the point mutation encoding lysine at residue 227 in the $\alpha 3$ domain.

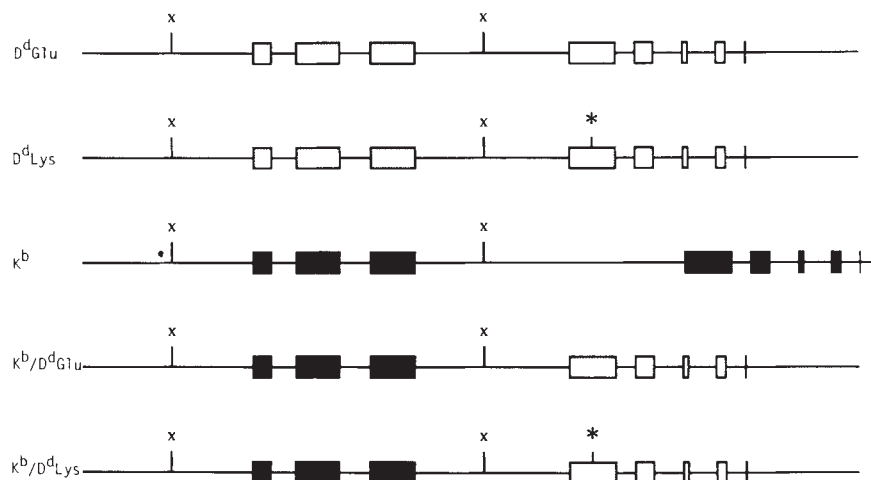
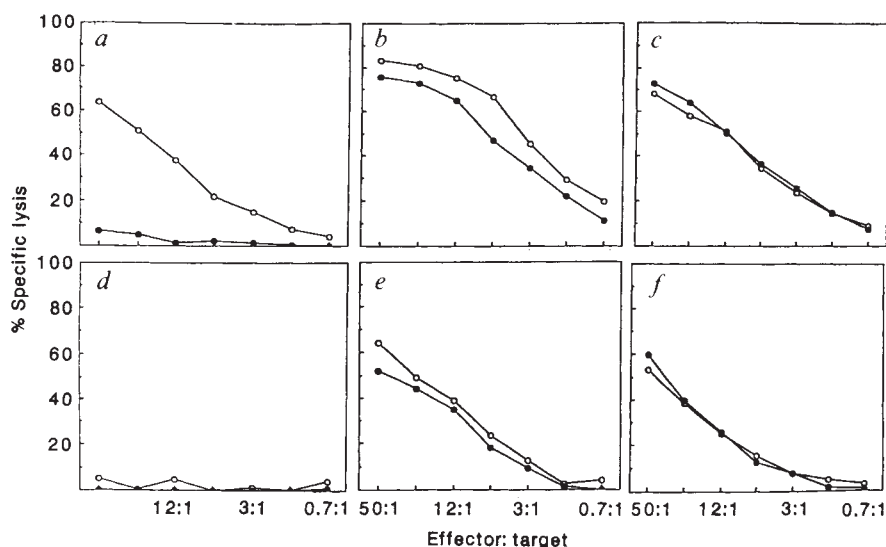


Fig. 3 Reactivity of primary and secondary anti-H-2K^b CTL for the R8.161 cells transfected with the H-2K^bD^d hybrid genes. The target cells were R8.161 H-2K^b/D^d Glu (a-c) and R8.161 H-2K^b/D^d Lys (d-f). The CTL populations were primary anti-H-2K^b (a and d), secondary anti-H-2K^b generated in the presence of anti-CD8 antibody (b and e) and primary anti-H-2^d (K^d, D^d, L^d) (c and f). The assay for CTL reactivity was performed either in the absence (○—○) or the presence (●—●) of an antibody to CD8.2 83-12-5. The inclusion of the anti H-2K^d/D^d CTL population (AKR anti B10.D2; panels c and f) demonstrated that both types of transfected cells were readily killed by CTL reactive with H-2^d molecules present on transfected and on non-transfected cells. The failure of the anti-CD8.2 antibody to inhibit these CTL, isolated from AKR mice which are CD8.1, established that the effect of this antibody on CTL populations isolated from B10.D2 was specific for the CD8 molecule.

Methods. The primary CTL were generated by a 5-day culture of B10.D2 spleen cells with irradiated (2,500 rad) B10 spleen cells, as previously described^{17,18}. 4×10^6 splenic responder cells were mixed with 4×10^6 stimulator cells in a final volume of 2 ml culture medium which contained selected fetal bovine serum (FBS) and 10% lectin-free concanavalin A-induced supernatant, and incubated at 37 °C for 5-6 days. The secondary CTL were generated by priming B10.D2 mice with irradiated (5,000 rad) EL4 tumour cells injected intraperitoneally one week before stimulation of the spleen cells *in vitro* in the presence of CD8.2 (83-12-5) antibody. The reactivity of CTL was assessed by ⁵¹Cr release, using 5×10^3 labelled target cells in a final volume of 200 μ l. Effector and target cells were incubated together in 96-well U-bottom plates for 4-6 h at 37 °C. Culture supernatants were collected and specific lysis determined as specific release (%): $100 \times (\text{experimental release}) - (\text{spontaneous release}) / (\text{maximum release}) - (\text{spontaneous release})$. The experimental release represents the release of ⁵¹Cr from target cells incubated with effector cells. The spontaneous lysis is the release in the absence of effector cells and the maximum release is that released by the target cells in the presence of 0.05 M HCl.



lation of the spleen cells *in vitro* for 5 days in the presence of antibody to CD8. These secondary CTL killed not only the H-2K^b/D^d Glu transfectant (Fig. 3b), but also the H-2K^b/D^d Lys transfectant (Fig. 3e). The level of lysis of the H-2K^b/D^d Lys transfectant was slightly less than that of the H-2K^b/D^d Glu transfectant, but a comparable difference was also found in the lysis of the transfectants by anti H-2^d CTL (Fig. 3c and f). Although the response of the secondary anti-H-2K^b CTL population was less susceptible than the primary anti-H-2K^b CTL to the addition of antibody to CD8, lysis of the H-2K^b/D^d Lys transfectant by the secondary CTL population was partially inhibited by the CD8 antibody (Fig. 3e). This effect could be due to the anti-CD8 antibody interfering either with the reaction of CD8 and other class I molecules on the surface of the transfected cells or with events such as receptor redistribution (capping) or signal transduction by the CD8 molecule.

Our interpretation of these findings is that substitution of lysine for glutamic acid at residue 227 destroys the determinant on the H-2 class I molecule that is recognized by the CD8 molecule on CTL. Although we think it is more likely that residue 227 contributes directly to the determinant recognized by CD8, we cannot rule out the possibility that substitution at residue 227 might indirectly alter the determinant bound by CD8. Another possibility is that the mutation at residue 227 could affect the conformation of the $\alpha 1/\alpha 2$ determinants, so that only high-affinity CD8-independent CTL can recognize the molecule. If there were such a major conformational change, however, then at least some of the anti H-2K^b $\alpha 1/\alpha 2$ monoclonal antibodies would probably have reacted differently with the H-2K^b/D^d Lys transfectant. As shown here, the antibodies reacted identically with both H-2K^b/D^d transfectants.

The three-dimensional structure of an HLA class I molecule has recently been determined^{12,13} and indicates that residues from the $\alpha 1$ and $\alpha 2$ domains could form a binding site for antigenic peptides between two α -helices. Residue 227 in the $\alpha 3$ domain is located on the surface of the molecule, is accessible to solvent and is at least 50 Å away from the $\alpha 1/\alpha 2$ peptide-

binding cleft (M. A. Saper and D. Wiley, personal communication), suggesting that substitution at position 227 is unlikely to affect CTL determinants in the $\alpha 1/\alpha 2$ domains dramatically. Another implication of our results is that, should CD8 bind to a conserved determinant which includes or is influenced by residue 227, then for CD8-dependent CTL, this interaction must be with the same class I molecule as is interacting with the T-cell receptor α/β complex. This is because the reactivity of primary CD8-dependent anti-H-2K^b CTL was abrogated by the mutation, even though the transfected cells expressed other class I molecules such as H-2K^d, H-2D^d and H-2L^d.

This work was supported by grants from the National Institutes of Health and a grant from the Lucille P. Markey Charitable trust. J.A.B. is a Gould Foundation Faculty Scholar. We thank John Hartley for expert technical assistance with the flow microfluorometry and Ms Rosina Passela for the preparation of this manuscript.

Received 2 September; accepted 11 November 1988.

- Swain, S. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7101-7105 (1981).
- MacDonald, H. R., Glasebrook, A. L., Bron, C., Kelsoe, A. & Cerrotini, J. C. *Immunol. Rev.* **68**, 89-115 (1982).
- Dembic, Z. *et al. Nature* **326**, 510-511 (1987).
- Gabert, J. *et al. Cell* **50**, 545-554 (1987).
- Fleischer, B., Schrezenmeier, H. & Wagner, H. *J. Immunol.* **136**, 1625-1628 (1986).
- van Seventer, G. A., van Lier, R. A. W., Spits, H., Ivanyi, P. & Melief, C. J. M. *Eur. J. Immunol.* **16**, 1363-1371 (1986).
- Yui, K., Hashimoto, Y., Wadsworth, S. & Greene, M. I. *J. exp. Med.* **166**, 1026-1040 (1987).
- Potter, T. A., Bluestone, J. A. & Rajan, T. V. *J. exp. Med.* **166**, 956-966 (1987).
- Potter, T. A., Palladino, M. A., Wilson, D. B. & Rajan, T. V. *J. exp. Med.* **158**, 1061-1076 (1983).
- Evans, G. A., Margulies, D. H., Shykind, B., Seidman, J. G. & Ozato, K. *Nature* **300**, 755-757 (1982).
- Allen, H., Wraith, D., Pala, P., Askonas, B. & Flavell, R. A. *Nature* **309**, 279-281 (1984).
- Bjorkman, P. J. *et al. Nature* **329**, 506-512 (1987).
- Bjorkman, P. J. *et al. Nature* **329**, 512-518 (1987).
- Margulies, D. H. *et al. J. Immunol.* **130**, 463-470 (1983).
- Potter, H., Weir, L. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7161-7165 (1984).
- Ajitkumar, P. *et al. Cell* **54**, 47-56 (1988).
- Levy, R. B., Ozato, K., Richard, J. C. & Bluestone, J. A. *J. Immunol.* **134**, 1342-1348 (1985).
- McCluskey, J. *et al. J. Immunol.* **137**, 3881-3890 (1986).

Duchenne muscular dystrophy gene product is not identical in muscle and brain

Uri Nudel, Dorit Zuk, Paz Einat, Elisha Zeelon*, Zehava Levy, Sara Neuman & David Yaffe

Department of Cell Biology, The Weizmann Institute of Science, Rehovot 76100, Israel

* Biotechnology General, Kiryat Weizmann, Ness Ziona, Israel

Duchenne muscular dystrophy (DMD) is an X-linked recessive disorder resulting in progressive degeneration of the muscle. It affects about 1 in 3,500 male children. Becker's muscular dystrophy is a less severe disease allelic to DMD. Some 30% of DMD patients suffer from various degrees of mental retardation (see ref. 1 for review). The giant DMD gene spans about 2,000 kilobases and codes for a 14-kilobase messenger RNA and a protein of molecular weight 427,000 (ref. 2). DMD mRNA is most abundant in skeletal and cardiac muscle and less so in smooth muscle³⁻⁵. We reported that the expression of the gene is developmentally regulated during the differentiation of primary muscle cultures and in myogenic cell lines in a way similar to the expression of muscle-specific genes such as myosin light chain 2 and skeletal muscle actin⁵. Similar results have been obtained with human primary myogenic cells⁶. Significant levels of DMD mRNA are found in brain tissue^{5,7}. Here we show that the transcript of the DMD gene and the amino terminal of the encoded protein differ in brain and muscle. The 5' ends of these mRNA species are derived from different exons. The results suggest that the two mRNA types are transcribed from different promoters.

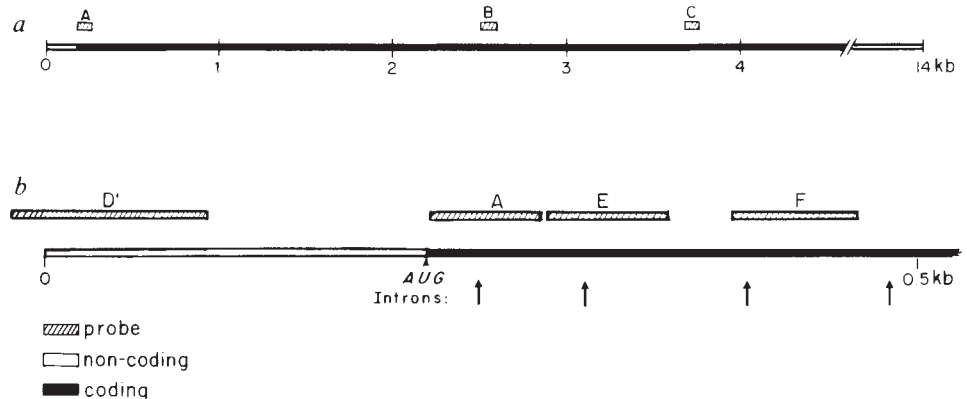
To investigate whether the DMD gene product is identical in muscle and brain, we prepared RNA probes complementary to various parts of the DMD mRNA. These RNA probes are described schematically in Fig. 1a. Total RNA preparations from rat skeletal muscle, cardiac muscle, aorta, brain and other tissues were assayed with these probes, using the RNase protection method⁸. Figure 2 shows how probes B and C, which are complementary to nucleotides 2,502-2,588 and 3,680-3,758 respectively of the reported mouse muscle DMD complementary DNA sequence^{4,9}, are protected by muscle and brain RNA, whereas probe A, which is complementary to the first 22 codons of the muscle mRNA, is protected by muscle mRNA and not by brain mRNA.

These results indicate that the 5' region of the muscle DMD mRNA contains sequences that are not present in the brain mRNA, so we constructed additional probes complementary to sequences in this region of the muscle DMD mRNA (Fig. 1b) and assayed them with muscle and brain RNA preparations. Probe D' was derived from a mouse genomic clone and is complementary to a sequence extending from the *Mbo*II site in the 5' noncoding region of the muscle type mRNA (nucleotide 80) into the 5' flanking region. A major 112-nucleotide-long fragment of this probe is protected by muscle RNA but not by brain RNA. These results were obtained with both rat and mouse mRNA. As the protected fragment is longer than expected (80 nucleotides), we suggest that the 5' end of the mRNA is located 32 nucleotides upstream from the position deduced from the published sequences of the cDNA clones^{4,9}. This conclusion is supported by primer extension experiments, and is in accordance with the position of a putative ATA box located 30 nucleotides upstream from the proposed major cap site (data not shown). Probes E and F cover codons 24-47 and 58-81 respectively and are protected by both muscle and brain RNA (Fig. 3; the results obtained with probe F are not shown).

On the basis of these results and of the proposed structure of the 5' region of the DMD gene (Fig. 1b), we concluded that the sequence of the exon that constitutes the 5' end of the muscle type DMD mRNA (deduced from the sequences of the muscle cDNA and genomic clones) is not represented in brain DMD mRNA. This exon contains the 5' untranslated region and the first 11 codons. The protection of probes E and F by both brain and muscle mRNA indicates that the brain DMD mRNA contains codons 24-47 and codons 58-81 of the muscle type mRNA.

To determine the exact point of divergence between the muscle and brain DMD mRNAs, we constructed a rat-brain cDNA library using an oligonucleotide primer complementary to nucleotides 554-582 of the muscle DMD mRNA. The library was screened with the probes described in the legend to Fig. 4 and a positive clone was purified, amplified and sequenced. Comparison of the rat brain and mouse muscle DMD cDNAs (Fig. 4) reveals an almost complete identity of the sequences starting at the 5' end of the muscle second exon (small sequence differences, which are all but one silent, are probably a result of the variability between rat and mouse). But sequences 5' to that point are completely different. Thus, the divergence point between the muscle and brain DMD mRNA types is at the junction between the first and second exons. This divergence also results in a difference between the amino terminal of the brain and muscle dystrophin proteins. The first eleven amino

Fig. 1 Schematic representation of the probes used in the RNase protection assay. *a*, The regions of DMD mRNA complementary to probes A, B and C. *b*, The 5' end of the muscle DMD mRNA and the probes used to analyse this region. The horizontal bar represents the DMD mRNA (blank, noncoding region; solid, coding region). The AUG indicates the putative translation initiation codon. The nucleotide numbering is according to the mouse cDNA sequences reported by Hoffman *et al.*⁴ and Koenig *et al.*⁹. Hatched boxes represent regions of mRNA complementary to the probes. Arrows (in *b*) indicate the position of introns as determined by sequencing of genomic and cDNA clones (ref. 2 and our unpublished data). Probe A is complementary to nucleotides 221-285 (the putative first 22 codons); probes B and C are complementary to nucleotides 2,502-2,588 and 3,680-3,758 respectively. Probe D' extends from the *Mbo*II site at nucleotide 80 into the 5' flanking region. Probes E and F are complementary to nucleotides 289-360 and 395-466 respectively. Probes A (pMD12), C (pMD51), E (pMD61) and F (pMD62) were constructed by introducing synthetic oligonucleotides into the vector gemini 3 (Promega). Probe B (pMD1), as described previously⁵, contains a rat genomic DNA fragment cloned in gemini 3. Probe D' (pMD13) contains a mouse genomic DNA fragment cloned in gemini 3. The RNA probes were synthesized using the T7 (probes A, C and D') or SP6 (probes B, E and F) RNA polymerases.



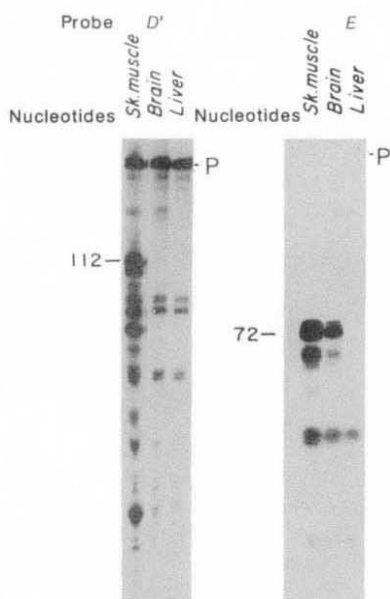
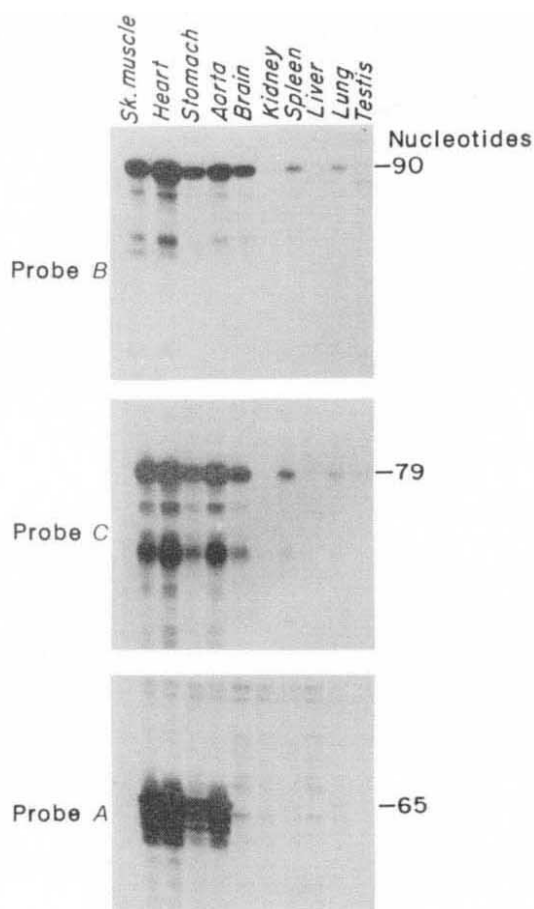


Fig. 3 Differential expression of sequences from the 5' end of the DMD mRNA in muscle and brain.

Methods. Total RNA (10 µg) from the rat tissues indicated were assayed with probes *D'* and *E* as described in Fig. 2, except that the RNase digestion after the hybridization with probe *D'* was for 1 h at 25 °C. P indicates the position of the undigested probes on the gel, and the numbers indicate the sizes of the main protected fragments. The smaller specific bands are probably caused by unstable regions in the RNA-RNA hybrids.

Fig. 2 DMD gene expression in various tissues.

Methods. Total RNA was prepared from the tissues indicated using the lithium chloride/urea extraction procedure¹¹. The integrity and quantity of all RNA samples were verified by electrophoresis on agarose/formaldehyde gels (not shown). All the RNA samples were extracted from rat tissue, except for testis RNA which was extracted from a mouse. As described previously⁵, there is a mismatch between the rat and the mouse sequences at the end of the sequence protecting probe *B*, giving a smaller protected band when mouse RNA is used, as can be seen for testis RNA in the middle panel. The RNA protection assay was as previously described⁵, except that hybridization was at 48 °C and 10 µg total RNA were used in each assay. Gels were fluorographed for four days after electrophoresis. *A*, *B* and *C* specify the probe used in each assay. As probes *A* and *C* contain vector sequences and probe *B* includes introns, the protected fragment is shorter than the labelled probe in all cases. The region of the gels containing the protected fragments is shown.

```
Brain: GACTTCTGGTTCCAGCAGCCGGCAGTAATAGAATGCTTTCAGGAAGATG 50
Muscle: TGTCAAGGCTGCTGAAGTTTATTTGGCTTCTCATCGTACCTAAGCCTCTCGT 150
ACAGAATCAGGAGAAAGATGCTGTTTTGCGCTATCTTGATTTGTTACAGC 100
GAGCAATAAACTGGGAGAACTTTTACCAAGATTTTATCCCTGCCTTG 200
AGCCAACTTATTGGCATGATGGAGTGACAGGAAAACAGCTGGCATGGAA 150
ATATATACTTTTTTCCAAATGCTTTGGTGGGAAGAAGTAGAGGACTGT 250
ATGAAAGAGAAGATGTTCAAAAGAAAACATTACAAAATGGATAAATGC 200
TATGAAAGAGAAGATGTTCAAAAGAAAACATTACAAAATGGATAAATGC 300
ACAATTTTCTAAGTTTGGAAAGCAGCACATAGACAACCTCTTCAGTGACC 250
ACAATTTTCTAAGTTTGGAAAGCAACACATAGACAACCTCTTCAGTGACC 350
TGCAGGATGGGAAACGCTCCTGGACCTCCTGGAAGGCCTGACAGGGCAA 300
TGCAGGATGGGAAACGCTCCTGACCTCCTGGAAGGCCTTACAGGGCAA 400
AAACTGCCAAAAGAAAAGGATCTACAGAGTTTCATGCCCTGAATAATGT 350
AAACTGCCAAAAGAAAAGGATCTACAGAGTTTCATGCCCTGAACAATGT 450
CAACAAGGCAGCTGCAGGTCTTACAGAAAATAATGTTGATTTAGTAATA 400
CAACAAGGCAGCTGCGGGTCTTACAGAAAATAATGTTGATTTAGTAATA 500
TAGGAAGCACTGACATAGTGGATGGAATCATAAACTCAC 440
TAGGAAGCACTGACATAGTGGATGGAATCATAAACTCAC 540
```

Fig. 4 Comparison of the 5' end region of the brain and muscle type DMD cDNAs. The rat brain DMD cDNA sequence was aligned with the mouse muscle DMD cDNA sequence. The triangles indicate the exon borders (ref. 2 and our unpublished data). The initiator ATGs of the two cDNAs are underlined. The brain type DMD mRNA contains five additional ATG triplets located upstream from the initiator ATG but they do not fit with the open reading frame of the coding sequences. These ATG triplets may be involved in the regulation of translation of the mRNA¹⁶. **Methods.** The brain cDNA library was constructed using polyadenylated RNA from rat brain and an oligonucleotide primer extending from nucleotide 554 to nucleotide 582 of the mouse muscle DMD cDNA⁹. The first and the second strands of the cDNA were synthesized using AMV reverse transcriptase and *E. coli* DNA polymerase respectively. After treatment with T4 DNA polymerase to produce blunt ends, ligation with *Eco*RI site-containing linkers, restriction with *Eco*RI and removal of the primer, the cDNA was cloned into the *Eco*RI site of λGT11. Approximately 50,000 phages were screened using two oligonucleotides homologous to the coding region of the mouse muscle DMD cDNA (nucleotides 318–360 and nucleotides 247–300). A phage that hybridized to both probes was purified and amplified. The 500-bp DNA insert was cloned into the *Eco*RI sites of M13mp19 and pUC19 and sequenced by the Sanger dideoxy termination method¹⁴ and by the Maxam and Gilbert chemical degradation method¹⁵.

acids (Met-Leu-Trp-Trp-Glu-Glu-Val-Glu-Asp-Cys-Tyr) of the muscle protein are replaced by three amino acids (Met-Glu-Asp) in brain dystrophin.

As the second exon is identical in both mRNA types, brain mRNA should also protect a 32-base pair (bp) fragment of probe A (Fig. 1), but we were unable to detect such a fragment in our assays (Fig. 2). This might be due to the very low G+C content (28%) of this region of the probe. The 5' half of the probe has a 45% G+C content, permitting stable hybridization when the entire probe is protected by muscle mRNA. As seen in Fig. 2, we can also detect DMD mRNA in spleen (5–10% of the amount in striated muscle), and smaller amounts are detected in the lung and testis. At the present level of sensitivity of the assay system, DMD mRNA was barely detectable in kidney and was not detectable in liver RNA preparations.

These results show that the DMD gene, which in myogenic cells is developmentally regulated, is also expressed in non-myogenic tissues. Detection of DMD mRNA and protein in non-muscle tissues has also been reported recently by Chelly *et al.*¹² and Hoffman *et al.*¹³ respectively. We have found that the DMD gene is also expressed in cultured brain cells (in collaboration with R. Simantov, unpublished results). This observation and the finding that the brain DMD mRNA is not identical to that of the muscle clearly demonstrates that the DMD RNA found in the brain tissue originated from brain cells and not from the smooth muscle of blood vessels. Experiments are currently being carried out to identify the DMD isoforms and the cell types that express the DMD gene in other non-muscle tissues.

Genetic considerations and Southern blot analysis have indicated the existence of only one DMD gene^{3,5}. Two mechanisms

could, separately or together, be responsible for the production of more than one mRNA from a single gene: (1) alternative splicing of an exon or exons from the gene transcript and (2) alternative initiation of transcription from different promoters reviewed in ref. 10). The finding that sequences of the 5' end of the muscle and brain DMD mRNAs are different indicates that two promoters with different tissue and developmental specificities could be involved in the initiation of transcription to produce the different mRNAs.

This work was supported by the US Muscular Dystrophy Association, the US-Israel Binational Science Foundation and the Leo and Julia Forchheimer Centre for Molecular Genetics at the Weizmann Institute of Science. We thank Dr P. Lonai for the mouse genomic DNA library and Ilana Liefer for technical assistance.

Received 29 February; revised 4 July; accepted 22 November 1988.

1. Moser, H. *Hum. Genet.* **66**, 17–40 (1984).
2. Koenig, M., Monaco, A. R. & Kunkel, L. M. *Cell* **53**, 219–228 (1988).
3. Monaco, A. P. *et al. Nature* **323**, 646–650 (1986).
4. Hoffman, E. P., Monaco, A. P., Feener, C. C. & Kunkel, L. M. *Science* **238**, 347–350 (1987).
5. Nudel, U., Robzyk, K. & Yaffe, D. *Nature* **331**, 635–638 (1988).
6. Lev, A. A., Feener, C. C., Kunkel, L. M. & Brown, R. H. Jr *J. biol. Chem.* **262**, 15817–15820 (1987).
7. Chamberlain, J. S. *et al. Science* **239**, 1416–1418 (1988).
8. Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035–7056 (1984).
9. Koenig, M. *et al. Cell* **50**, 509–517 (1987).
10. Breitbart, R. E., Andreadis, A. & Nadal-Ginard, B. *A. Rev. Biochem.* **56**, 467–495 (1987).
11. Auffray, C., Nageotte, R., Chambraud, B. & Rougeon, F. *Nucleic Acids Res.* **8**, 1231–1241 (1980).
12. Chelly, J., Kaplan, J.-C., Maire, P., Gautron, S. & Kahn, A. *Nature* **333**, 858–860 (1988).
13. Hoffman, E. P., Hudecki, M. S., Rosenberg, P. A., Pollina, C. M. & Kunkel, L. M. *Neuron* **1**, 411–420 (1988).
14. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
15. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
16. Mueller, P. & Hinnebusch, A. G. *Cell* **45**, 201–207 (1986).

Effects of the tumour promoter okadaic acid on intracellular protein phosphorylation and metabolism

T. A. J. Haystead, A. T. R. Sim, D. Carling,
R. C. Honnor*, Y. Tsukitani†, P. Cohen
& D. G. Hardie

MRC Protein Phosphorylation Group, Biochemistry Department,
The University, Dundee DD1 4HN, UK

* Department of Biochemistry and Genetics, Medical School,
University of Newcastle upon Tyne, NE2 4HH, UK

† Fujisawa Pharmaceutical Company, Tokyo, Japan

Okadaic acid is a polyether derivative of 38-carbon fatty acid¹, and is implicated as the causative agent of diarrhetic shellfish poisoning². It is a potent tumour promoter that is not an activator of protein kinase C (ref. 3), but is a powerful inhibitor of protein phosphatases-1 and -2A (PP1 and PP2A) *in vitro*^{4,5}. We report here that okadaic acid rapidly stimulates protein phosphorylation in intact cells, and behaves like a specific protein phosphatase inhibitor in a variety of metabolic processes. Our results indicate that PP1 and PP2A are the dominant protein phosphatases acting on a wide range of phosphoproteins *in vivo*. We also find that okadaic acid mimics the effect of insulin on glucose transport in adipocytes, which suggests that this process is stimulated by a serine/threonine phosphorylation event.

The first insight into the mechanism of action of okadaic acid came from the finding that it enhanced the contraction of skinned smooth muscle fibres, indicating that myosin P-light chain phosphorylation was increased⁶. This appeared to be mediated by inhibition of protein phosphatase(s) as myosin light chain kinase activity was unaffected and a phosphatase from smooth muscle with activity towards isolated myosin P-light chains was inhibited by okadaic acid⁷. The toxin was recently shown to be

a potent inhibitor of PP1 and PP2A (refs 4 and 5), which have homologous catalytic subunits⁸ and make up two of the four cytosolic serine/threonine-specific protein phosphatase catalytic subunits in mammalian cells^{9–11}. PP2B (a calcium/calmodulin-dependent enzyme) was inhibited with much lower potency, whereas PP2C (a magnesium-dependent enzyme) was unaffected^{4,5}.

Okadaic acid produces marked increases in the phosphorylation states of many proteins in adipocytes and hepatocytes, as judged by increased ³²P-labelling (Fig. 1), without changing the specific radioactivity of intracellular ATP or the ATP:ADP ratio (data not shown). Proteins whose phosphorylation increased included polypeptides that comigrated with acetyl-CoA carboxylase and ATP-citrate lyase in adipocyte and hepatocyte cytosol, pyruvate kinase and 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PFK2/FBPase) in hepatocyte cytosol, and glycogen phosphorylase and glycogen synthase in the hepatocyte glycogen fraction. The prominent protein of relative molecular mass 32,000 in the hepatic microsomal fraction whose phosphorylation increased markedly in response to okadaic acid is likely to be ribosomal protein S6. Okadaic acid treatment (15 min) stimulated the ³²P-labelling of total trichloroacetic acid-insoluble protein in hepatocyte cytosol 2.5-fold, with a half-maximal effect at 200 nM and maximal effect at 1 µM.

The protein kinases which phosphorylate the above-mentioned proteins *in vivo* include cyclic AMP-dependent protein kinase¹², the AMP-activated protein kinase¹³, phosphorylase kinase¹⁴, glycogen synthase kinase-3 (ref. 15) and casein kinase-2 (refs 16 and 17). Using glycogen synthase as substrate, okadaic acid at concentrations as high as 5 µM failed to affect the activity of any of these protein kinases, or casein kinase-1 (data not shown). The purified catalytic subunits of PP1 and PP2A were potently inhibited by okadaic acid with 50% inhibitory concentration values of 20 nM and 0.2 nM respectively, which are even lower than those reported previously^{4,5}. The PP2A catalytic subunit was completely inhibited by 1 nM

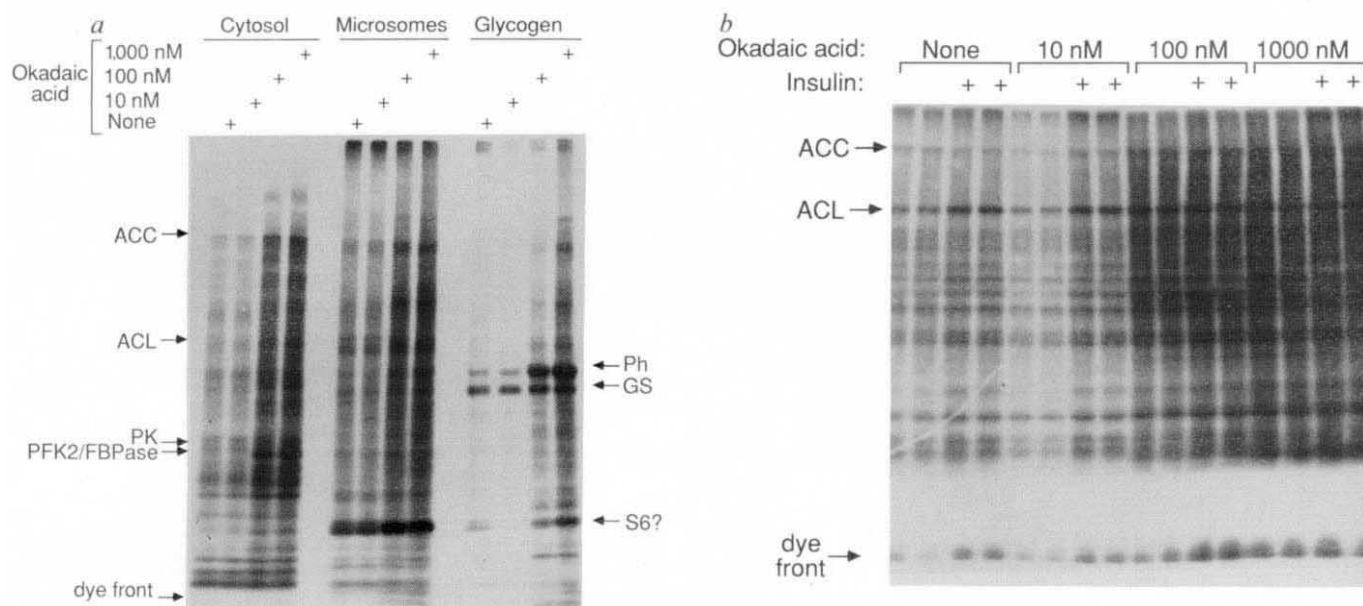


Fig. 1 Effects of okadaic acid on protein phosphorylation in isolated hepatocytes (*a*) and adipocytes (*b*) from fed rats. For adipocytes, incubations were carried out in duplicate in the presence and absence of 0.9 nM insulin. Cells were isolated and prelabelled with [32 P]phosphate for 60 min^{20,21} before addition of okadaic acid or insulin for 15 min. Okadaic acid was added in 10% dimethylsulphoxide to give a final concentration of dimethylsulphoxide of 0.25% (v/v) in all incubations. Cells were collected and homogenized, and post-mitochondrial supernatants were prepared by centrifugation (10,000g for 15 min at 4°C). For hepatocytes only, the post-mitochondrial supernatant was separated into cytosol, glycogen and microsomal fractions by centrifugation (100,000g for 60 min at 4°C). The supernatants were removed (cytosol fraction), and the upper microsomal pellet was resuspended in homogenization buffer separately from the lower glycogen pellet. Fractions were analysed at equal protein loadings by electrophoresis in 5–15% gradient polyacrylamide gels. Gels were dried and autoradiographed at -70°C using Amersham Hyperfilm-MP in Kodak X-omatic intensifying cassettes. The migration of marker proteins is shown, using the abbreviations: ACC, acetyl CoA carboxylase; ACL, ATP-citrate lyase; PK, pyruvate kinase; PFK2/FBPase, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase; Ph, glycogen phosphorylase; GS, glycogen synthase; S6, ribosomal protein S6. The specific radioactivity of cellular ATP and ATP/ADP ratios were determined by HPLC (ref. 21).

okadaic acid, a concentration comparable with that of the phosphatase subunit in the assay.

As a number of the proteins in adipocytes and hepatocytes whose phosphorylation was increased by okadaic acid were enzymes of glucose and lipid metabolism, we examined the effects of okadaic acid on the flux through these pathways. Okadaic acid mimicked the action of glucagon in increasing the rate of glucose output from hepatocytes, as well as the conversion of lactate to glucose (Fig. 2), as would be expected if the toxin increased phosphorylation of glycogen phosphorylase, glycogen synthase, pyruvate kinase and PFK2/FBPase. Okadaic acid stimulated basal lipolysis in adipocytes 2.6-fold (Fig. 3*a*), which is consistent with increased phosphorylation of hormone-sensitive lipase. A larger stimulation of lipolysis (10-fold) was observed in the presence of the adenosine agonist PIA, which causes inhibition of adenylate cyclase and reduces basal cAMP-dependent protein kinase activity and lipolysis¹⁷ (Fig. 3*c*). Okadaic acid also completely prevented insulin from antagonizing the β -adrenergic activation of lipolysis (Fig. 3*b*).

Phosphorylation is known to inhibit acetyl CoA carboxylase, the rate-limiting enzyme in fatty acid biosynthesis¹⁸, and we found that the activity of this enzyme was greatly decreased in extracts from okadaic acid-treated adipocytes (Fig. 3*d*). Acetyl CoA carboxylase activity itself was unaffected by inclusion of okadaic acid (1 μ M) in the assays. There was a good correlation between inactivation of acetyl-CoA carboxylase and inhibition of insulin-stimulated fatty acid synthesis, as measured by incorporation of radioactivity from [14 C]-labelled acetate into total lipid (Fig. 3*d*). Okadaic acid also inhibited the low rate of fatty acid synthesis observed in the absence of insulin (not shown).

Although okadaic acid antagonized the actions of insulin on fatty acid synthesis and lipolysis (Fig. 3), it mimicked the action

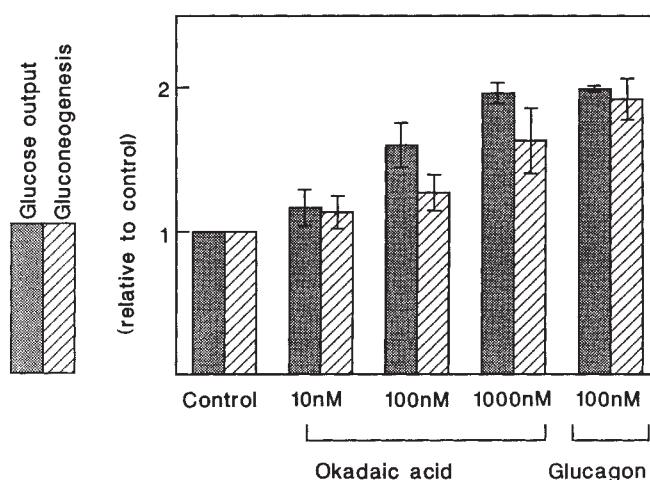


Fig. 2 Effects of okadaic acid on glucose output and gluconeogenesis in isolated hepatocytes from fed rats. Hepatocytes were isolated in medium containing 10 mM glucose and preincubated for 10 min with okadaic acid as in Fig. 1, except that the medium contained unlabelled phosphate. Gluconeogenesis was measured by the incorporation of radioactivity from [14 C]lactate into glucose²³. Glucose output was measured using glucose oxidase (Sigma diagnostic kit 510-A); for these studies only, cells were collected by centrifugation and resuspended in medium without glucose immediately before addition of okadaic acid.

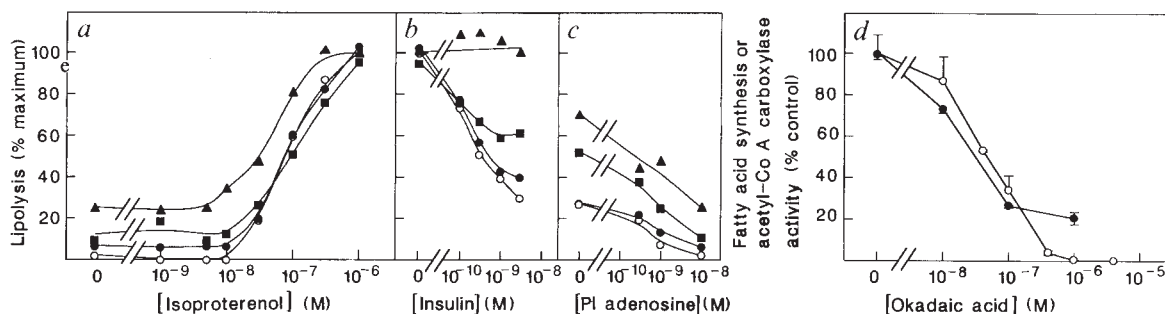


Fig. 3 Effects of okadaic acid treatment on metabolic parameters of isolated adipocytes. *a-c*, Effects of incubation for 25 min without okadaic acid (open circles), or with okadaic acid at 100 nM (filled circles), 300 nM (squares) or 1 μ M (triangles) on lipolysis measured as glycerol release¹⁷. The adenosine agonist PIA (*N*⁶-[*R*-(-)-1-methyl-2-phenethyl]adenosine, 100 nM) was present in *a*, and isoproterenol (1 μ M) and PIA (100 nM) were present in *b*. *d*, Incorporation of radioactivity from [¹⁴C]acetate into toluene extractable-lipid²⁴ (open circles) and acetyl CoA carboxylase activity (closed circles) measured at 10 mM citrate in the post-mitochondrial supernatant fraction²⁰.

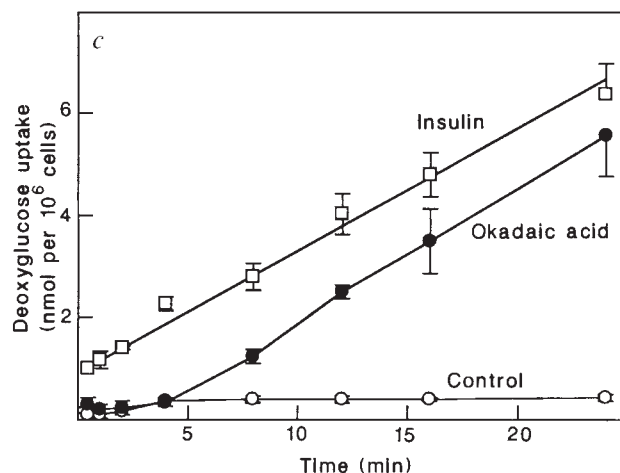
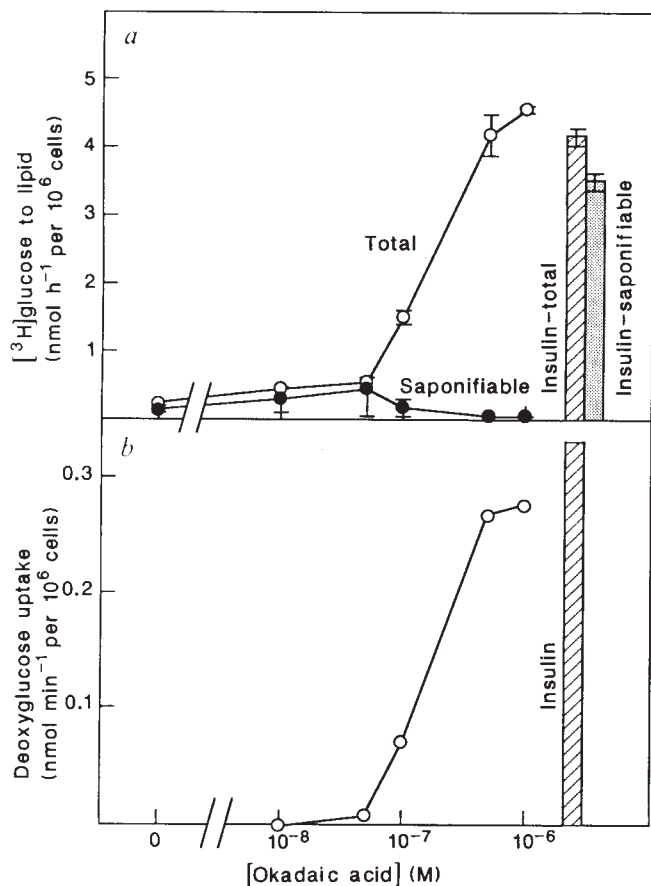


Fig. 4 Effects of okadaic acid and insulin on conversion of glucose into lipid and on 2-deoxyglucose uptake in adipocytes. *a*, Effect of okadaic acid on incorporation of radioactivity from [³H]glucose into total toluene-extractable lipid (open circles) or saponifiable lipid (filled circles)²⁰. The effects of a 10 min incubation with 0.9 nM insulin (total lipid, cross-hatched bar; saponifiable lipid, shaded bar) are also shown for comparison. *b*, Effect of okadaic acid on uptake of [¹⁴C]-labelled 2-deoxyglucose into cells. The effect of a 10 min incubation with 0.9 nM insulin is shown by the cross-hatched bar. For both *a* and *b*, cells were preincubated with okadaic acid for 10 min before the measurement. *c*, Time course for uptake of [¹⁴C]-labelled 2-deoxyglucose into cells treated in the absence of okadaic acid or insulin (open circles), or in the presence of 0.9 nM insulin (open squares) or 1 μ M okadaic acid (filled circles). Deoxyglucose uptake was measured by incubating cells with 0.5 mM [¹⁴C] 2-deoxyglucose (0.2 μ Ci ml⁻¹). Reactions were stopped by adding 200 μ l aliquots of cells to 50 μ M cytochalasin B (50 μ l) and 50 μ M phloretin in 0.25 $\times$ 4 cm polythene microtubes. 100 μ l silicone oil (0.97 g ml⁻¹) was layered on top of the cells, and the tubes were then spun at 7,000g for 5 s in a bench centrifuge. The top half of the tubes containing the cells was cut off with a scalpel and dropped into 1 ml scintillation fluid (LKB Optiphase II) for counting. Control incubations showed that cytochalasin B and phloretin completely inhibited 2-deoxyglucose uptake in response to okadaic acid.

of insulin in dramatically enhancing the incorporation of radioactivity from [³H]glucose into total lipid (Fig. 4a). But whereas 84% of the radioactivity incorporated into lipid was saponifiable (as fatty acid) in the case of insulin, <5% was saponifiable after treatment with okadaic acid. The simplest explanation for these results is that okadaic acid stimulated glucose transport while activating lipolysis and inhibiting fatty acid synthesis, so that most of the increased radioactivity would be glycerol-incorporated during triglyceride turnover. Insulin, however, stimulates glucose transport while inhibiting lipolysis and activating fatty acid synthesis, and radioactivity incorporated from glucose into triglyceride in this case would largely be present in fatty acid. To measure the effect of okadaic acid on glucose transport more directly, we studied the uptake of 2-deoxyglucose into adipocytes. After a lag of about 5 min, okadaic acid greatly stimulated uptake. Insulin produced the same effect, but without a significant lag (Fig. 4).

Our results strongly support the view that the effects of okadaic acid on cell function are explained by its ability to inhibit PP1 and/or PP2A. All effects on intact cells were maximal at 1 μ M, which is similar to the *in vivo* concentration of PP1 and PP2A (ref. 11). None of the eight protein kinases tested to date (our data and refs 3 and 7) were affected by okadaic acid. Okadaic acid at 1 μ M scarcely affects PP2B activity⁵, and its specificity

is further emphasized by its failure to inhibit PP2C (ref. 4), pyruvate dehydrogenase phosphatase (T.A.J.H., unpublished results), protein tyrosine phosphatases, acid and alkaline phosphatases, and inositol trisphosphatase⁴. Furthermore, because okadaic acid stimulates glycogenolysis and gluconeogenesis in hepatocytes, and lipolysis and the conversion of glucose to glyceride-glycerol in adipocytes (Figs 2-4), and does not affect ATP or ADP levels, none of the enzymes in these metabolic pathways appears to be inhibited non-specifically by this compound.

Assays in cell-free extracts have shown that PP1, PP2A and PP2C account for virtually all the phosphatase activity in skeletal muscle and liver towards some 20 phosphoproteins that regulate several metabolic pathways and muscle contraction⁹⁻¹¹. This study now strongly suggests that PP1 and PP2A, rather than PP2C, are the dominant phosphatase catalytic subunits acting on a wide range of phosphoproteins *in vivo*. The toxin is therefore a powerful probe for the investigation of the role of protein phosphorylation and of PP1/PP2A in any cellular event. It has already been shown to increase the calcium current in isolated myocytes⁵, suggesting that the slow inward calcium channel is activated by phosphorylation in cardiac cells, and inactivated by PP1 and/or PP2A.

Although most of the effects of okadaic acid on protein phosphorylation and metabolism could have been predicted, the stimulation of glucose transport was unexpected, and the results are evidence that this process is activated by a protein serine/threonine phosphorylation event. The failure to observe phosphorylation of the glucose transporter in response to insulin¹⁹, suggests that an associated regulatory protein may be the target for phosphorylation, rather than the transporter itself.

Because PP1 and PP2A are likely to be the chief enzymes that reverse the actions of protein kinase C, it is not surprising that a phosphatase inhibitor which can enter intact cells should be as potent a tumour promoter as substances that activate protein kinase C. Tumour promotion presumably stems from increased phosphorylation of one or more proteins that are substrates for protein kinase C and dephosphorylated by PP1/PP2A. The effects of okadaic acid on protein phosphorylation in isolated cells (Fig. 1) emphasize what a powerful tumour-suppressing effect PP1 and PP2A must have on normal cells.

The studies at Dundee were supported by the Medical Research Council, the Wellcome Trust and the British Heart Foundation. We thank Simon Pilakis for gifts of purified pyruvate kinase and PFK2/FBPase, and Richard Denton for useful discussions.

Received 12 October; accepted 11 November 1988.

1. Tachibana, K. *et al.* *J. Am. Chem. Soc.* **103**, 2469-2471 (1981).
2. Murata, M., Shimatani, M., Sugitani, H., Oshima, Y. & Yasumoto, T. *Bull. Japan Soc. scient. Fish.* **48**, 549-552 (1982).
3. Suganuma, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 1768-1777 (1988).
4. Bialojan, C. & Takai, A. *Biochem. J.* **256**, 283-290 (1988).
5. Hescheler, J., Mieskes, G., Ruegg, J. C., Takai, A. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **412**, 248-252 (1988).
6. Ozaki, H., Kohama, K., Nonomura, Y., Shibata, S. & Karaki, H. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **335**, 356-358 (1987).
7. Takai, A., Bialojan, C., Troschka, M. & Ruegg, J. C. *FEBS Lett.* **217**, 81-84 (1987).
8. Berndt, N. *et al.* *FEBS Lett.* **223**, 340-346 (1987).
9. Ingebritsen, T. S. & Cohen, P. *Science* **221**, 331-338 (1983).
10. Cohen, P. *Proc. R. Soc. B* **234**, 115-144 (1988).
11. Cohen, P. *A. Rev. Biochem.* (in the press).
12. Pilakis, S. J., El-Maghrabi, M. R. & Claus, J. H. A. *Rev. Biochem.* **57**, 755-783 (1988).
13. Sim, A. T. R. & Hardie, D. G. *FEBS Lett.* **233**, 294-298 (1988).
14. Cohen, P. *Molec. Asp. Cell. Reg. Vol. 5* (eds Cohen, P. & Klee, C. B.) 123-144 (Elsevier Biomedical, Amsterdam, 1988).
15. Cohen, P. in *The Enzymes* 3rd edn Vol. 17 (eds Boyer, P. D. & Krebs, E. G.), 361-392 (Academic, Orlando, 1987).
16. Haystead, T. A. J., Campbell, D. G. & Hardie, D. G. *Eur. J. Biochem.* **175**, 347-354 (1988).
17. Honnor, R. C., Dhillon, G. S. & Londos, C. *J. Biol. Chem.* **260**, 15122-15129 (1985).
18. Munday, M. R., Campbell, D. G., Carling, D. & Hardie, D. G. *Eur. J. Biochem.* **175**, 331-338 (1988).
19. Gibbs, E. M., Allard, W. J. & Lienhard, G. E. *J. Biol. Chem.* **261**, 16597-16603 (1986).
20. Haystead, T. A. J. & Hardie, D. G. *Biochem. J.* **234**, 279-284 (1986).
21. Holland, R., Witters, L. A. & Hardie, D. G. *Eur. J. Biochem.* **140**, 325-333 (1984).
22. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
23. Exton, J. H. & Park, C. R. *J. Biol. Chem.* **244**, 1424-1433 (1968).
24. Haystead, T. A. J. & Hardie, D. G. *Eur. J. Biochem.* **175**, 339-345 (1988).

The pseudoautosomal boundary in man is defined by an *Alu* repeat sequence inserted on the Y chromosome

Nathan A. Ellis, Paul J. Goodfellow, Barbara Pym, Matthijs Smith, Mark Palmer, Anna-Maria Frischauf & Peter N. Goodfellow

Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK

The Y chromosome, which in man determines the male sex (reviewed in ref. 1), is composed of two functionally distinct regions. The pseudoautosomal region is shared between the X and Y chromosome and is probably required for the correct segregation of the sex chromosomes during male meiosis². The second region includes the sex-determining gene(s), the presence of which is necessary for the development of testes. The two regions have contrasting genetic properties: the pseudoautosomal region recombines between the X and Y chromosome³; the Y-specific region must avoid recombination otherwise the chromosomal basis of sex-determination breaks down¹. The pseudoautosomal region is bounded at the distal end by the telomere and at the proximal end by X- and Y-specific DNA. We have found that the proximal boundary was formed by the insertion of an *Alu* sequence on the Y chromosome early in the primate lineage. Proximal to the *Alu* insertion there is a small region where similarity between the X and Y chromosomes is reduced and which is no longer subject to recombination.

Previous studies provided us with a defined region for the location of the pseudoautosomal boundary⁴. The *MIC2* gene is the closest marker known distal to the pseudoautosomal boundary and the probe (3'Rsa) is derived from the 3'-most exon of the *MIC2* gene. *DYS104* is the most distal Y-specific sequence that has been described⁴. To isolate the pseudoautosomal boundary we cloned 110 kilobases (kb) of DNA between the sequences defined by 3'Rsa and *DYS104* (Fig. 1). The Y-derived cosmid cAMF31 contains sequences that cross the pseudoautosomal boundary because it includes both Y-specific (H2.1) and pseudoautosomal sequences (E1.8). This places the boundary between 30 kb and 45 kb away from the sequences defined by 3'Rsa.

Two probes (Hf0.2 and RsY0.55) from cosmid cAMF31 were used to define the pseudoautosomal boundary: both probes are located on a Y-specific *SacI* fragment of 3.3 kb (*SacI*Y3.3; Fig. 2a). Hybridization of Hf0.2 to *EcoRI* and *HaeIII* digests of DNA from females, from males and from somatic cell hybrids having the X chromosome and the Y chromosome as their only human contribution, reveals fragments of 1.8 kb and of 0.6 kb respectively (Fig. 2b, lanes 1-8). When Hf0.2 is hybridized to the same DNAs after *SacI* digestion, fragments of 3.3 kb and 4.5 kb appear in male DNA but only one fragment of 4.5 kb is evident in female DNA (Fig. 2b, lanes 13 and 14). The 3.3-kb fragment is specific to the Y chromosome (lane 16) and the 4.5-kb fragment is specific to the X chromosome (lane 15). Similarly, Hf0.2 recognizes 2.4-kb and 2.8-kb *RsaI* fragments in male DNA and only a 2.8-kb fragment in female DNA (lanes 9 and 10). The 2.4-kb fragment comes from the Y (lane 12), and the 2.8-kb fragment comes from the X (lane 11). Hence, the pseudoautosomal Hf0.2 probe maps Y-specific and X-specific restriction sites. The pseudoautosomal boundary lies between the *EcoRI* and *RsaI* sites proximal to the sequences defined by Hf0.2.

The results in Fig. 2c confirm this conclusion. Hybridization of the RsY0.55 probe to the *SacI*-digested DNAs presented in Fig. 2b reveals the 3.3-kb fragment in male and 'Y-only' DNA, but not in female and 'X-only' DNA. DNA from 64 females

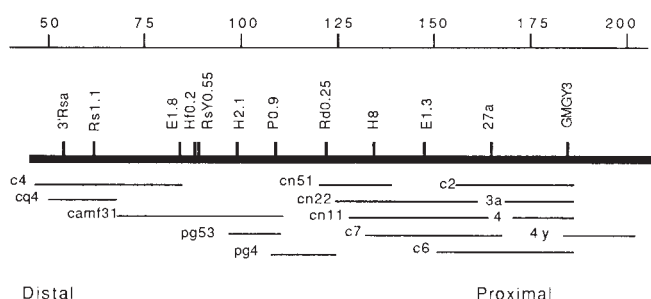


Fig. 1 Map of the region cloned in the chromosome walk. If the 5' CpG-rich region of *MIC2* is 0 kb, we set 3'Rsa at 55 kb on the long-range restriction map on the basis of partial cloning and Southern blotting of the *MIC2* gene^{4,8}. Unique probes were isolated at each stage of the walk and their chromosomal origin determined by hybridization to gel transfers of DNA from female (WT49), male (PGF), 'X-only' somatic cell hybrid (C12D) and 'Y-only' somatic cell hybrid (853). Probe designations and clone sizes are shown. Cosmid clones are preceded by 'c'; all others are lambda clones. Probe 27a defines *DYS104* and GMGY3 defines *DYS13*. Probes H8, Rd0.25 and E1.8 need competition with sheared human placenta DNA⁹.

Methods. Three cosmid libraries (one was donated by Peter Little and constructed with male DNA; the two others were constructed with male DNA and 853 DNA) and four lambda libraries (HL100 was constructed with male DNA and purchased from Clontech, and the other three were constructed with 853 DNA) were used in screening. Libraries were made by ligating genomic DNA, either partially digested with *Mbo*I or completely digested with *Eco*RI or *Hind*III, into appropriately digested pcos8 or λ DASH (Stratagene) DNA, followed by packaging and transduction into P2392, ED8767 or NM621 *Escherichia coli* strains. Hybridization was in 6 $\times$ SSC, 5 $\times$ Denhardt's solution, 0.5% SDS, 0.25 mg ml⁻¹ sheared and denatured DNA from salmon testes, and 2 ng ml⁻¹ labelled probe. Probes were isolated on NA45 paper and labelled with ³²P by oligonucleotide reaction¹⁰. For hybridization with cosmid libraries, 0.1 mg ml⁻¹ *Hae*III-digested and denatured vector DNA was added to prevent non-specific binding. The filters were washed twice at room temperature in 2 $\times$ SSC and 0.2% SDS, then 3–5 times for 15 min in 0.2 $\times$ SSC and 0.2% SDS at 65 $^{\circ}$ C, mounted and then exposed to XAR-5 with Cronex lightning-plus intensifier screens overnight.

and 57 males was hybridized with the RsY0.55 probe: all the male samples, but none of the female samples, reacted with the probe. Therefore, the proximal pseudoautosomal region rarely, if ever, includes the sequences defined by RsY0.55 located on the Y chromosome.

X-specific sequences adjacent to the boundary were isolated by screening a lambda library with a probe derived from SacIY3.3. Overlapping clones λ X31 and λ X21 contain the pseudoautosomal 1.8-kb *Eco*RI fragment and the X-specific 4.5-kb *Sac*I fragment. The 4.5-kb *Sac*I fragment (SacIX4.5), which contains the X-side of the pseudoautosomal boundary, was subcloned and mapped using restriction enzymes (Fig. 3). The maps of SacIX4.5 and SacIY3.3 are nearly identical as far as the *Pst*I sites in the middle of each clone, but beyond this they are dissimilar. The pseudoautosomal boundary seems to lie in the region of ~50-base pair (bp) between the *Msp*I and *Hha*I sites on the Y chromosome (Fig. 3).

The 0.9-kb *Eco*RI/*Rsa*I fragment from SacIY3.3 (RsE0.9) and the 0.8-kb *Pst*I fragment from SacIX4.5 (P0.8) were subcloned (Fig. 3) and sequenced (Fig. 4a and b). As predicted by the restriction map, the boundary is located near the *Hha*I site. The two sequences diverge dramatically 245 bp proximal to the pseudoautosomal *Eco*RI site. Only two bases differ between the Y and X chromosomes at locations 201 bp and 205 bp from the *Eco*RI site. At the boundary on the Y chromosome, a 303-bp *Alu* repeat is inserted (Fig. 4b) and the first

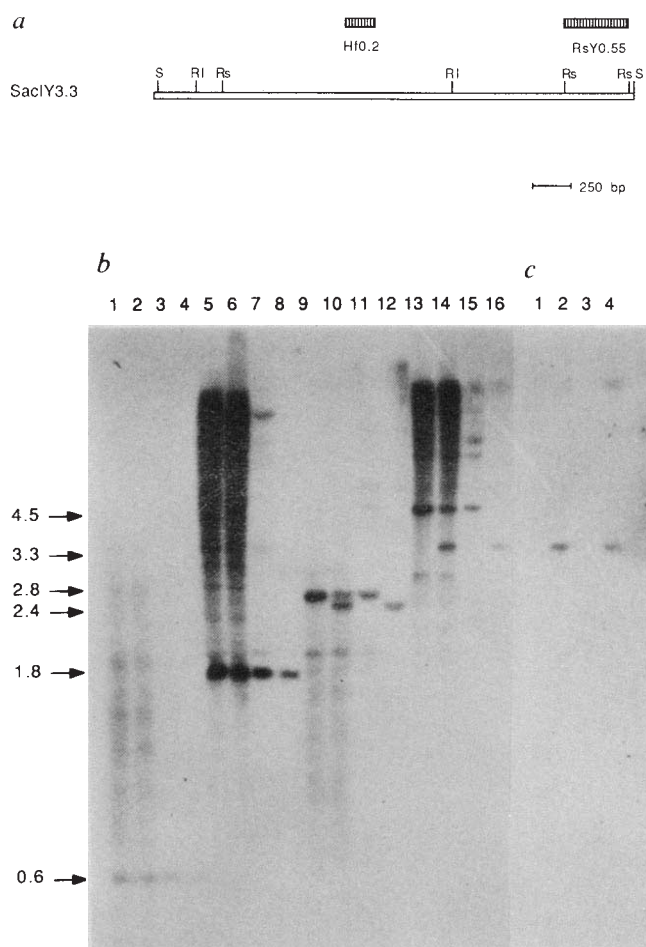
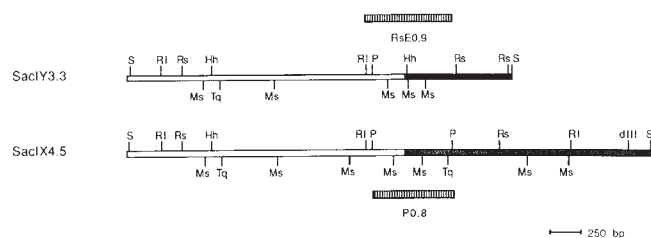


Fig. 2 Mapping the pseudoautosomal boundary with the Hf0.2 probe. **a**, Restriction map of SacIY3.3 and the two probes derived from it. Hf0.2 defines pseudoautosomal sequences and RsY0.55 defines Y-specific sequences. The 3.3-kb *Sac*I fragment from cAMF31 was subcloned into the *Sma*I site of pUC8. Hf0.2 and RsY0.55 were also subcloned into the *Sma*I site of pUC8. S, *Sac*I; RI, *Eco*RI; Rs, *Rsa*I. **b**, The Hf0.2 probe was hybridized to female, male, 'X-only' hybrid and 'Y-only' hybrid DNAs: WT49 (lanes 1, 5, 9 and 13), PGF (lanes 2, 6, 10 and 14), C12D (lanes 3, 7, 11 and 15) and 853 (lanes 4, 8, 12 and 16). Lanes 1–4 show DNA digested with *Hae*III; lanes 5–8, with *Eco*RI; lanes 9–12, with *Rsa*I and lanes 13–16, with *Sac*I. **c**, RsY0.55 was hybridized to the filter used in **b** after stripping it first. Lanes 1–4 are *Sac*I-digested DNA from female, male, 'X-only' hybrid and 'Y-only' hybrid. **Methods.** Each DNA (10 μ g) was digested according to the supplier's instructions, electrophoresed through a 0.7% agarose gel and transferred on to GeneScreen Plus (NEN). Hybridization, post-hybridization washing and autoradiography have been previously described¹¹.

base pair of the *Alu* insertion is the first base-pair difference between the X and Y chromosomes. On the other side of the *Alu* insertion, homology between the X and Y chromosomes resumes and extends for 225 bp. The X and Y show 77% homology in this region. Beyond this point the sequences cannot be aligned by the IFIND:ALIGN (Intelligenetics) program. Figure 4c shows a schematic representation of the boundary region. Two regions of homology are separated by an *Alu* insertion, and proximal to these the sequences are non-homologous.

We propose that the pseudoautosomal boundary is precisely located at the point where the *Alu* sequence is inserted on the Y chromosome. Recombination has prevented sequence divergence in the region distal to the *Alu* (*Alu*-distal region). If the



◀ **Fig. 3** A restriction map of the pseudoautosomal boundary. The lambda clones from which the X-specific 4.5-kb *SacI* fragment is derived were isolated from a flow-sorted X-chromosome library (Los Alamos, ID code LA0XNL01). The fragment was subcloned into the *SacI* site of Bluescript[®] (Stratagene). RsE0.9 was subcloned into *EcoRI*/*SmaI*-digested M13mp18 DNA and P0.8 into *PstI*-digested pUC8 DNA. One *MspI* site present in the *SacI*X4.5 is missing in the *SacI*Y3.3. This site is associated with a restriction-fragment length polymorphism (data not shown).

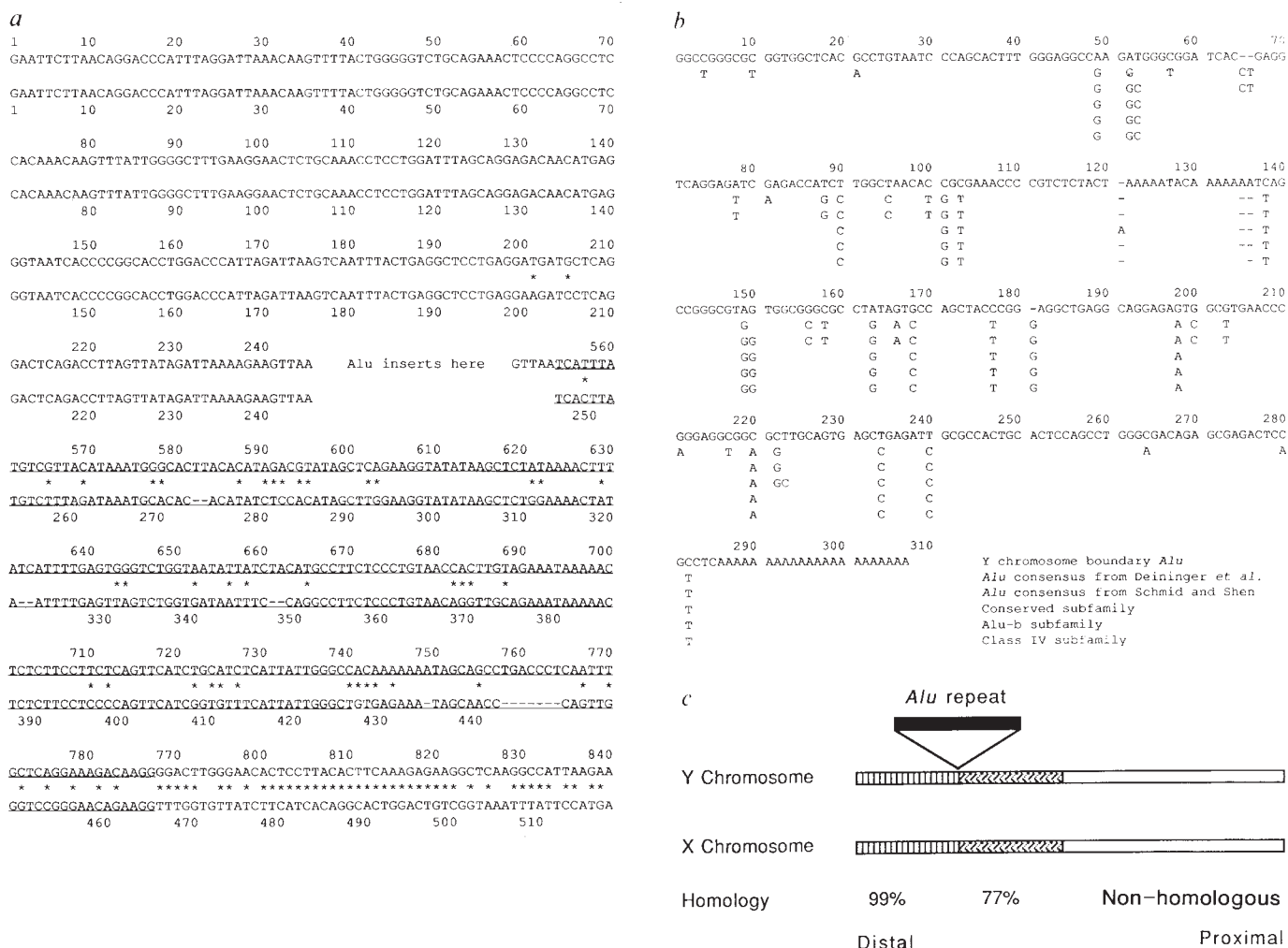


Fig. 4 DNA sequence of the pseudoautosomal boundary. Subclones were derived from the RsE0.9 and P0.8 in either pUC8 or Bluescript[®] and sequenced in double-stranded DNA by the dideoxy chain-termination method^{12,13}. Oligonucleotides were used to complete the project. The sequence from the *EcoRI* site to the distal *PstI* site on the X chromosome was derived from the appropriate 1.8-kb *EcoRI* subclone of SacIX4.5. *a*, The DNA sequence before and following the *Alu* insert on the X and Y chromosome is compared by removing the *Alu*. The *Alu*-proximal region is shown underlined, mismatches are starred and gaps are dashed. Note that multiple base changes frequently occur in the *Alu*-proximal region and that the gaps are all on the X chromosome. The complete sequence of the 947-bp RsE0.9 fragment and 771 bp from *EcoRI* to *PstI* from SacIX4.5 is available on request. *b*, The *Alu* insert is compared with the consensus sequences from Deininger *et al.*¹⁴ and from Schmid and Shen¹⁵, and with the subfamily consensus sequences from Willard *et al.*¹⁶, Jurka and Smith¹⁷, and Britten *et al.*¹⁸. The *Alu* insertion appears to be a retroposition¹⁹ because it ends with a string of 22 adenine nucleotides and has 5-bp direct repeats at each end (GTAA). The *Alu* belongs to the conserved subfamily and has diverged by 6–7% from this group. *c*, Schematic drawing of the pseudoautosomal boundary shows the aligned *Alu* proximal and *Alu* distal regions with percentage homology represented as hatched lines, the *Alu* insert as a solid line, and X- and Y-specific regions as open lines.

mutation rate exceeded the probability of crossover between the mutation and the boundary, mutations would accumulate near the boundary. The pseudoautosomal region undergoes recombination at an average rate of 1% (10^{-2}) per 5×10^4 bp, therefore the recombination rate is 2×10^{-7} per meiosis per base pair³. If we make the parsimonious assumption that there is one meiosis per generation, and three generations every 100 years, we get a recombination rate of 6×10^{-9} per bp per year. Given the rate

of sequence evolution is $\sim 1.5 \times 10^{-9}$ mutations per site per year⁵, then the rate of recombination slightly exceeds the mutation rate at the boundary. The *Alu*-distal region shows only two base differences out of 245 bp, which is consistent with our analysis; these two differences may be pseudoautosomal polymorphisms.

The fact that the region proximal to the *Alu* (*Alu*-proximal region) has diverged 23% implies that recombination has not maintained sequence similarity in this region. If recombination

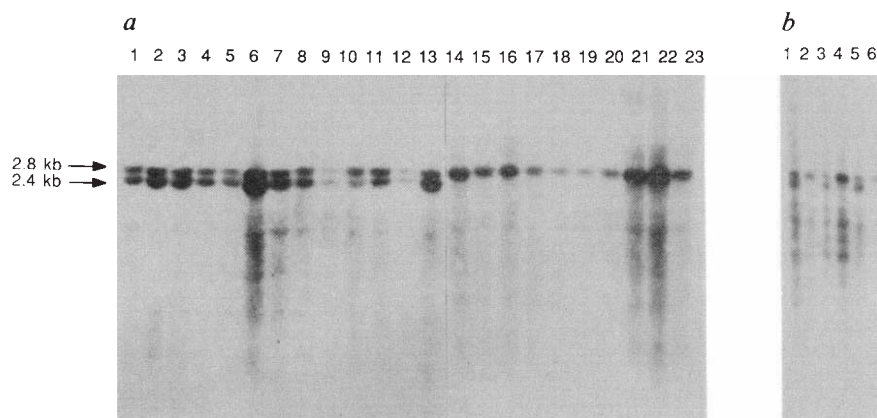


Fig. 5 Hybridization of the Hf0.2 probe to *RsaI*-digested DNA from males and females. Female DNA shows a single 2.8-kb X-derived fragment; male DNA gives a 2.8-kb X-derived fragment and a 2.4-kb Y-derived fragment. **a**, DNAs from human males are in lanes 1–13 and DNAs from females in lanes 14–23. A total of 50 males, 25 females, an 'X-only' and a 'Y-only' hybrid was tested. Gel transfers were hybridized with RsY0.55 to test for Y-specific DNA. **b**, Male (odd-numbered lanes) and female (even-numbered lanes) DNAs from human (lanes 1 and 2), chimpanzee (lanes 3 and 4) and gorilla (lanes 5 and 6).

were occurring in the *Alu*-proximal region, then the *Alu* insertion would be present in females as well as in males. To test this hypothesis, we hybridized the Hf0.2 probe to *RsaI*-digested DNA from males and females. Hf0.2 detects *RsaI* fragments of 2.4 kb and 2.8 kb which cross the pseudoautosomal boundary. If recombination is occurring within the *Alu*-proximal region, the *Alu* is transferred to the X and increases the size of the X-specific *RsaI* fragment to 3.1 kb, decreasing the Y-specific *RsaI* fragment to 2.1 kb. We have already shown that sequences defined by the RsY0.55 are not found in females, but if recombination were to occur within or proximal to this locus, then hybridization with Hf0.2 would reveal a 2.4-kb *RsaI* fragment in female DNA. We have tested 50 males and 25 females for this effect (see Fig. 5a for a subset of these results). In all cases, males had two fragments of 2.4 kb and 2.8 kb and females had only a 2.8-kb fragment. Moreover, in DNA from male and female gorillas and chimpanzees, we detected *RsaI* fragments of identical size (Fig. 5b). We conclude that *Alu* insertion is of ancient origin and that recombination is not occurring in the *Alu*-proximal region. The base differences in the *Alu*-proximal region are uniformly distributed throughout the 225 bp, which suggests that divergence began for the whole region at once. This allows us to estimate the age of the human pseudoautosomal boundary. The common ancestor of the 225-bp region on the X and Y chromosome differs by 11.5% ($23\% \div 2$) from that found today, which dates the boundary at ~45 million years old⁶. Unless other structural rearrangements have occurred, the previous boundary was 225 bp proximal to the *Alu* insertion and was also abrupt (Fig. 4a), as sequences proximal to the *Alu*-proximal region are non-homologous.

It is clear that the boundary has been created by the juxtaposition of nonhomologous sequences. Local maintenance of the boundary appears to be due, at least in part, to suppression of recombination caused by the nonhomologous sequences. Previously we suggested that the boundary could be limited by sex chromosome-specific functional genes or regulatory elements⁷. DNA sequencing and searching of databases has failed to provide evidence for such sequences close to the boundary on the X or the Y chromosome.

In summary, we have cloned and sequenced the pseudoautosomal boundary in man. The present-day boundary is defined by the insertion of a 303-bp *Alu* repeat sequence in the Y chromosome. Sequences immediately adjacent to the *Alu* insertion are identical, implying recombination occurs up to the *Alu*. The *Alu* insertion itself is Y-specific. Therefore the boundary is an abrupt change located at the *Alu*. Proximal to the *Alu* is a 225-bp region of X–Y homology which has diverged by 23%, suggesting the existence of an earlier boundary located 225 bp proximal to the *Alu*. We suggest that the absence of sequence homology was enough to create the pseudoautosomal boundary. Studies in primates should definitively date the *Alu* insertion.

We thank Peter Little for a cosmid library, Nigel Spurr for

gel transfers of DNA from males and females, Ian Goldsmith of the ICRF's oligonucleotide synthesis group for oligonucleotides, and Kim Richardson for secretarial assistance.

Received 30 September; accepted 23 November 1988.

- Goodfellow, P. N. & Darling, S. M. *Development* **102**, 251–258 (1988).
- Burgoyne, P. S. *Hum. Genet.* **61**, 85–90 (1982).
- Petit, C., Leveilliers, J. & Weissenbach, J. *EMBO J.* **7**, 2369–2376 (1988).
- Pritchard, C. A., Goodfellow, P. J. & Goodfellow, P. N. *Nature* **328**, 273–275 (1987).
- Miyamoto, M. M., Slightom, J. L. & Goodman, M. *Science* **238**, 369–373 (1987).
- Koop, B. F., Goodman, M., Xu, P., Chan, K. & Slightom, J. L. *Nature* **319**, 234–238 (1986).
- Bengtsson, B. O. & Goodfellow, P. N. *A. Hum. Genet.* **51**, 57–63 (1987).
- Goodfellow, P. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 5605–5609 (1988).
- Litt, M. & White, R. L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6206–6210 (1985).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6–13 (1983).
- Goodfellow, P. J., Darling, S. M., Thomas, N. S. & Goodfellow, P. N. *Science* **234**, 740–743 (1986).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4767–4771 (1987).
- Deininger, P. L., Jolly, D. J., Rubin, C. M., Friedmann, T. & Schmid, C. W. *J. molec. Biol.* **151**, 17–33 (1981).
- Schmid, C. W. & Shen, C. K. J. in *Molecular Evolutionary Genetics* (ed. MacIntyre, R.) 323–358 (Plenum, New York, 1986).
- Willard, C., Nguyen, H. T. & Schmid, C. W. *J. molec. Evol.* **26**, 180–186 (1987).
- Jurka, J. & Smith, T. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4775–4778 (1988).
- Britten, R. J., Baron, W. F., Stout, D. B. & Davidson, E. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4770–4774 (1988).
- Weiner, A. M., Deininger, P. L. & Efstratiadis, A. A. *Rev. Biochem.* **55**, 631–661 (1986).

Mitochondrial splicing requires a protein from a novel helicase family

Bertrand Séraphin*, Michel Simon, Annick Boulet & Gérard Faye

Institut Curie-Biologie, Bat. 110, Centre Universitaire, 91405 Orsay Cedex, France

Proteins involved in mitochondrial splicing but encoded by nuclear genes have been characterized in *Saccharomyces* and *Neurospora*^{1–5}. The role in splicing of these proteins is largely unknown. Here we report that mutations in the nuclear gene *MSS116* directly affect the splicing of several introns of the cytochrome *b* (*cob*) and cytochrome *c* oxidase subunit I (*cox1*) primary transcripts. This implies that the MSS116 protein (pMSS116) is an important component of the mitochondrial splicing machinery. The sequence of the cloned *MSS116* gene shows that its protein product is homologous to the translation eIF-4A factor and the human nuclear protein p68. We show further that these proteins share several conserved amino-acid blocks with DNA helicases and related proteins. This suggests that pMSS116 has an RNA helicase activity. RNA helicases may be involved in many different processes including translation and splicing.

* Present address: Department of Biology, Brandeis University, Waltham, Massachusetts 02254, USA

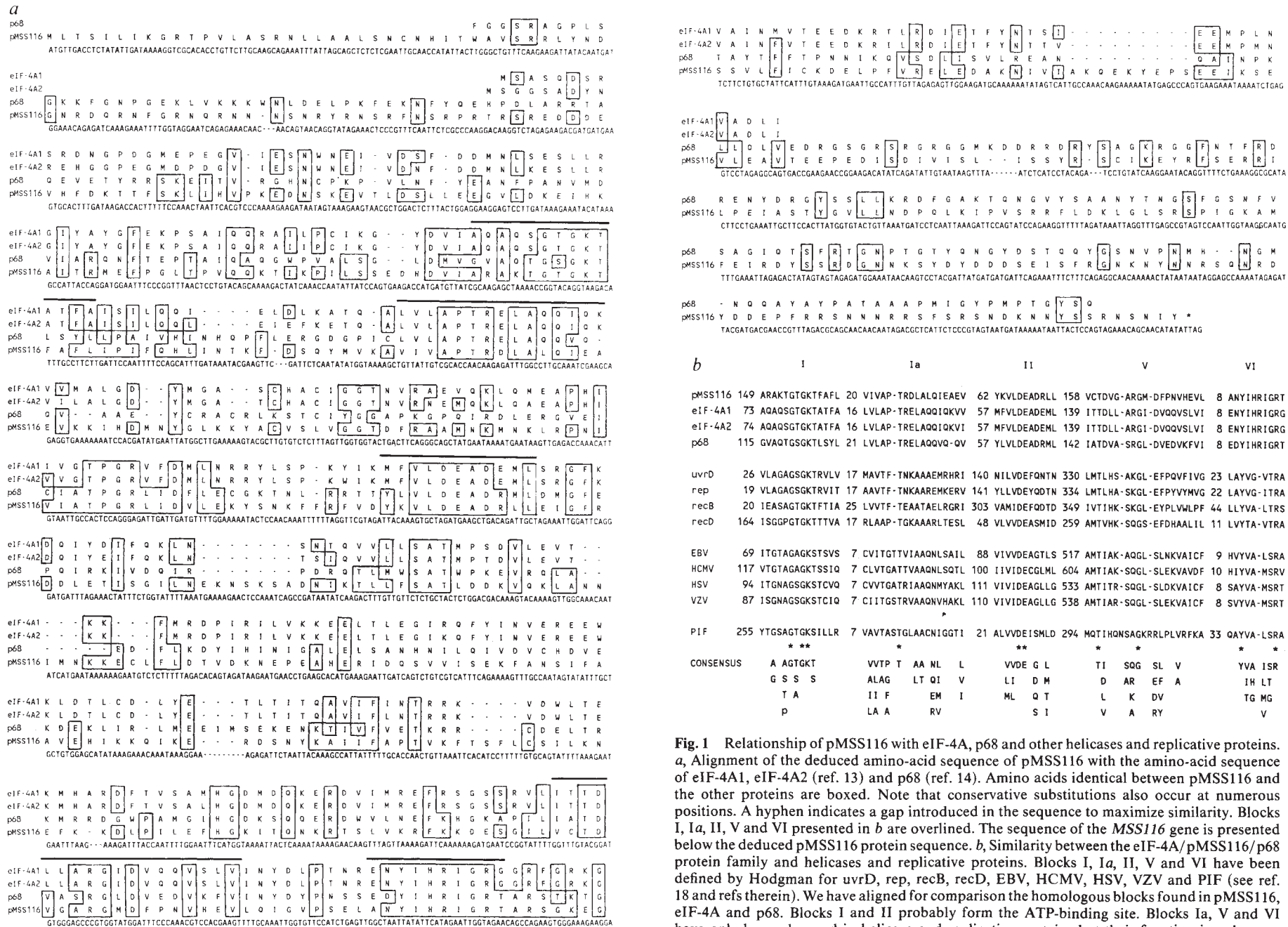


Fig. 1 Relationship of pMSS116 with eIF-4A, p68 and other helicases and replicative proteins. **a**, Alignment of the deduced amino-acid sequence of pMSS116 with the amino-acid sequence of eIF-4A1, eIF-4A2 (ref. 13) and p68 (ref. 14). Amino acids identical between pMSS116 and the other proteins are boxed. Note that conservative substitutions also occur at numerous positions. A hyphen indicates a gap introduced in the sequence to maximize similarity. Blocks I, Ia, II, V and VI presented in **b** are overlined. The sequence of the *MSS116* gene is presented below the deduced pMSS116 protein sequence. **b**, Similarity between the eIF-4A/pMSS116/p68 protein family and helicases and replicative proteins. Blocks I, Ia, II, V and VI have been defined by Hodgman for uvrD, rep, recB, recD, EBV, HCMV, HSV, VZV and PIF (see ref. 18 and refs therein). We have aligned for comparison the homologous blocks found in pMSS116, eIF-4A and p68. Blocks I and II probably form the ATP-binding site. Blocks Ia, V and VI have only been observed in helicase and replicative proteins but their function is unknown. Asterisks indicate amino acids conserved in all sequences. The consensus was derived by selecting positions where four or fewer amino acids occur.

We recently described a genetic screen for yeast nuclear mutants which specifically affect mitochondrial splicing⁶. Among the selected mutants, *petA179* (renamed *mss116-2* here) is allelic to *mss116-1* which was shown, by Northern screening analysis, to impede the splicing of *cox1* transcript⁷. Northern experiments showed that mutation *mss116-2* has a drastic effect on the processing of primary transcripts (pre-mRNA) in a strain harbouring five introns in the *cob* gene and seven in the *cox1* gene (see Fig. 2 in ref. 6). A block in the splicing of one or several mitochondrial introns gives rise to a respiratory deficient phenotype. To find out which of the 13 known introns is dependent upon a wild-type *MSS116* protein for splicing, mitochondrial genomes containing various sets of introns were introduced by cytoduction^{6,8} into strains carrying either of the two mutant alleles. Lack of growth of the resulting cytoductants on glycerol (that is, their inability to respire) indicates which intron(s) requires the wild-type *MSS116* gene product for splicing. The results of these experiments, described in Table 1, indicate that excision of introns a11 as well as a15 α and/or a15 β of *cox1* and b11, b12 and/or b13 of *cob* pre-mRNA depends upon the activity of pMSS116.

To gain further insight into the *MSS116* gene function, we cloned the wild-type allele by complementation in yeast and localized it within a 3-kb *HindIII* fragment. Its sequence shows a long reading frame coding for a putative protein of 664 amino acids (Fig. 1a). Gene disruption experiments⁹ confirmed that we had cloned the *MSS116* gene (data not shown). Interestingly, introduction of mitochondria devoid of the 13 mitochondrial introns into a strain carrying a disrupted *MSS116* gene (allele *mss116-3*) does not restore respiratory competence. This suggests that pMSS116 has multiple functions. An alternative explanation is that some unidentified intron may exist in yeast mitochondria.

When we searched data bases for proteins homologous to the pMSS116 amino-acid sequence, we found a remarkable homology with the murine translation factor eIF4A (ref. 10) (Fig. 1a). To show that pMSS116 is directly involved in the splicing of *cob* and *cox1* pre-mRNA, and is not a translation initiation factor, we analysed the incorporation of [³⁵S]methionine into mitochondrial proteins *in vivo*¹¹. Results obtained from strains bearing a disrupted *MSS116* gene are shown in Fig. 2. It appears that in a strain carrying the 13 known mitochondrial introns (Fig. 2, lane 1) the *cob* and *cox1* proteins are absent, whereas the other major mitochondrial proteins are expressed. But we have not shown whether both the *cox3* and ATPase 6 proteins are synthesized. When these mitochondria are

exchanged by cytoduction for those of GF134-6D (see Table 1), protein *cob* is indeed translated and a low, though detectable, level of protein *cox1* is synthesized (Fig. 2, lane 3). These results show that pMSS116 is not a factor affecting general translation initiation *per se*. But our preliminary analyses of mitochondrial translation in *mss116* strain does not rule out the possibility that pMSS116 is an initiation factor specifically required for efficient translation of the *cox1*, *cox3* or ATPase 6 mRNA (Fig. 2, lane 3). An alternative possibility is that pMSS116 may be in some way necessary for the synthesis of some intron-encoded proteins (maturases). Northern experiments revealed that in the mitochondrial M12-54 context the splicing efficiency of intron b11 is greatly reduced by the *mss116-2* mutation (data not shown). However b11 does not require mitochondrial-encoded maturases for its excision¹². Thus, at least for intron b11, pMSS116 does not intervene in splicing via the synthesis of maturases. These data suggest that pMSS116 is directly involved in the excision of mitochondrial introns.

Recently a second murine initiation factor gene was characterized¹³. Its amino-acid sequence and that of a human protein p68 (ref. 14) is compared with pMSS116 and eIF4A in Fig. 1a. Alignment of the four proteins together displays a remarkable conservation of several long amino-acids blocks covering most of the eIF4A sequences (although the middle part is slightly less conserved). Pairwise comparison gives 22.1%, 26.9% and 26.4% identical residues between pMSS116 and p68, eIF4A1 and eIF4A2, respectively. Conservative substitutions have occurred at several other positions. The extreme N-terminal part of pMSS116 has no counterpart in any of the other three proteins. This region is rich in serine, threonine and basic amino acids, and is devoid of acidic residues. These are the characteristics of mitochondrial pre-sequences¹⁵ and support the idea that pMSS116 is imported into the mitochondrion. The eIF4A polypeptide is involved in the melting or unwinding of RNA in the presence of ATP^{16,17}. The human p68 polypeptide is supposed to have a DNA helicase activity *in vitro*¹⁴. This suggests that pMSS116 may also possess a helicase activity.

This suggestion is further strengthened by comparison of the protein sequences with recently reported consensus sequences found in replicative proteins¹⁸. These proteins include the *uvrD* and *rep* helicases as well as *recB*, *UL5* and related proteins which are likely to possess helicase activity (see ref. 18 and references therein). Out of the seven domains described by Hodgman in these helicase and replicative proteins, five are clearly identifiable in pMSS116, p68 and eIF4A (Fig. 1b). Blocks I and II probably form part of the ATP binding site, whereas

Table 1 Mitochondrial intron contents of the *cob* and *cox1* genes, and the respiratory competence of strains carrying the *mss116-1* or *mss116-2* alleles

Origin of mitochondria	Mitochondrial intron contents													Respiratory competence	
	ω	b1	b2	<i>cob</i>			<i>cox1</i>							allele <i>mss116-1</i>	allele <i>mss116-2</i>
(a) M12-54	—	+	+	+	+	+	+	+	+	+	+	+	+	—	—
(b) GF204-1D	+	—	—	—	+	+	—	+	+	+	+	+	+	—	ND
(c) D273-10B/A	+	—	—	—	+	+	+	+	+	+	—	—	+	—	—
(d) D273-10B/G1/356-R6	+	—	—	—	+	+	—	—	+	+	—	—	+	+	+
(e) GF134-6D	+	—	—	—	+	+	—	—	+	+	—	—	—	ND	+
(f) UVA32	—	—	—	—	+	+	—	+	+	—	—	—	+	+	+
(g) GF167-7B	—	—	—	—	—	—	—	—	—	—	—	—	—	+	+

Strains indicated in the column 'origin of mitochondria' (GF204-1D excepted) are described in ref. 6. Mitochondria of these seven strains were introduced by cytoduction, via strain K5/2 (ref. 6), either in strain BS7-76-2/2 carrying allele *mss116-1* or in strain D273-M3A/A179/F11 carrying allele *mss116-2*. These three latter strains had their mitochondrial genome deleted (ρ^0) (ref. 6). Mitochondrial intron contents: ω is the 21S ribosomal RNA intron, b1 to b5 are the cytochrome b introns (b1) and a1 to a5 γ are the cytochrome oxidase subunit 1 introns (a1); + or — indicates that the intron is present or deleted. Introns b1, a1, a2 and a5 γ belong to group II, the others to group I (compare with ref. 20 and references therein). Respiratory competence was tested on glycerol plates: + or —, growth or absence of growth; ND, not determined. Assuming that a block of the splicing of a15 α and/or a15 β does not hamper the splicing of any of the five *cob* introns, we can reach the following conclusions. Mutation *mss116-1* in strain harbouring the mitochondria of D273-10B/A impedes the splicing of *cox1* pre-mRNA (ref. 7), whereas the cytochrome b protein is synthesized^{11,25}; comparison of lines a and c implies that *MSS116* is necessary for the splicing of intron b11, b12 and/or b13. By comparing lines c, d and f and knowing that in the D273-10B/A mitochondrial context the splicing of intron a11 is severely deficient (data not shown) we can conclude that *MSS116* is necessary for the splicing of intron a11 and/or a15 α and/or a15 β . Comparison of lines c and b leads to the conclusion that *MSS116* is necessary for the splicing of a11 as well as a15 α and/or a15 β .

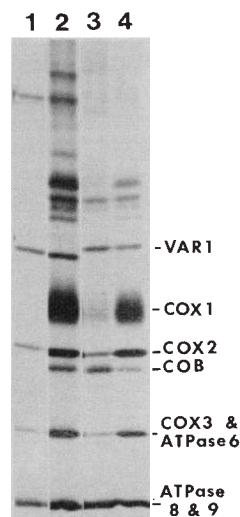


Fig. 2 ^{35}S -labelled mitochondrial translation products of strains, GRF18/B601B (lane 1), GRF18 (lane 2), BS102-1A (lane 3) and BS102-1B (lane 4). Cells were grown in galactose or raffinose complete media. They were labelled with [^{35}S] methionine in the presence of cycloheximide¹¹. Radiolabelled mitochondrial translation products were prepared and analysed by SDS-PAGE²⁴. Strains GRF18/B601B and BS102-1A bear a disrupted *MSS116* gene (allele *mss116-3*), the others a wild-type *MSS116* gene. Strains GRF18/B601B and GRF18 carry the 13 known mitochondrial introns, strains BS102-1A and BS102-1B have the mitochondria of GF134-6D (compare Table 1). The positions of the mitochondrially synthesized polypeptides are indicated: VAR1, var1 ribosomal protein; cox1, cox2 and cox3, subunits I, II and III of cytochrome c oxidase; COB, cytochrome b; ATPase 6, ATPase 8 and ATPase 9, subunits 6, 8 and 9 of mitochondrial oligomycin-sensitive ATPase.

blocks Ia, V and VI have only been observed in helicases and related proteins (Fig. 1b). These data indicate that the helicase/replicative protein super-family may be extended to include the more divergent eIF4A/p68/pMSS116 subfamily and reinforce the assumption that pMSS116 is an RNA helicase.

Several models may explain how an RNA helicase affects mitochondrial splicing. For example, folding of the nascent transcripts may give rise to several secondary structures in various proportions and only some of these structures will be compatible with a self-splicing activity. In support of this, only a fraction of the RNA present in an *in vitro* self-splicing reaction is active and the transition to the reactive state is the rate-limiting step¹⁹. An RNA helicase activity, by allowing a cycle of folding and unfolding to take place, will help each molecule reach the active configuration driving the reaction towards splicing of all pre-mRNA molecules. A differential propensity of various introns to non-functional folding will be reflected by different requirements for a helicase.

Another hypothesis is that changes in secondary and/or tertiary structures, which are likely to occur during splicing of mitochondrial introns, may be catalysed by a structure and/or sequence-specific helicase. Thus a folding pattern required early in the splicing process is melted later by a helicase allowing a new structure to form. During nuclear splicing, which is somewhat related to group II intron splicing²⁰, the U4 and U6 small nuclear RNAs follow a cycle of association and disassociation^{21,22}, and it has been suggested that a helicase may have a role in melting the two molecules apart^{21,23}.

We thank H. Fukuhara for encouragement and support, C. Wilson and R. Chanet for a critical reading of the manuscript.

Received 21 September; accepted 14 November 1988.

- Hill, J., McGraw, P. & Tzagoloff, A. *J. biol. Chem.* **260**, 3235-3238 (1985).
- Kreike, J., Schulze, M., Pillar, T., Kört, A. & Rödel, G. *Curr. Genet.* **11**, 185-191 (1986).

- Séraphin, B., Simon, M. & Faye, G. *EMBO J.* **7**, 1455-1464 (1988).
- Akins, R. & Lambowitz, A. *Cell* **50**, 331-345 (1987).
- Herbert, C., Labouesse, M., Dujardin, G. & Slonimski, P. *EMBO J.* **7**, 473-483 (1988).
- Séraphin, B., Boulet, A., Simon, M. & Faye, G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6810-6814 (1987).
- Faye, G. & Simon, M. in *Mitochondria 1983* (eds Schweyen, R., Wolf, K. & Kaudewitz, F. 433-439 (de Gruyter, Berlin, 1983).
- Conde, J. & Fink, G. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3651-3655 (1976).
- Rothstein, R. *Meth. Enzym.* **101**, 202-211 (1983).
- Nielsen, P., McMaster, G. & Traschel, H. *Nucleic Acids Res.* **13**, 6867-6880 (1985).
- McKee, E., McEwen, J. & Poyton, R. *J. biol. Chem.* **259**, 9332-9338 (1984).
- Hensgens, L. *et al. J. molec. Biol.* **164**, 35-58 (1983).
- Nielsen, P. & Traschel, H. *EMBO J.* **7**, 2097-2105 (1988).
- Ford, M., Anton, I. & Lane, D. *Nature* **332**, 736-738 (1988).
- Roise, D. & Schatz, G. *J. biol. Chem.* **263**, 4509-4511 (1988).
- Ray, B. *et al. J. biol. Chem.* **260**, 7651-7658 (1985).
- Abramson, R. *et al. J. biol. Chem.* **262**, 3826-3832 (1987).
- Hodgman, T. *Nature* **333**, 22-23 (1988).
- Zaug, A., Kent, J. & Cech, T. *Biochemistry* **24**, 6211-6218 (1985).
- Cech, T. *Cell* **44**, 207-210 (1986).
- Chang, S. & Abelson, J. *Gene Dev.* **1**, 1014-1027 (1987).
- Lamond, A., Konarska, M., Grabowski, P. & Sharp, P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 411-415 (1988).
- Brow, D. & Guthrie, C. *Nature* **334**, 213-218 (1988).
- Douglas, M., Finkelstein, D. & Butow, R. *Meth. Enzym.* **56**, 58-66 (1979).
- Tzagoloff, A., Akai, A. & Needleman, R. *J. biol. Chem.* **250**, 8228-8235 (1975).

The gene for the U6 small nuclear RNA in fission yeast has an intron

Tokio Tani & Yasumi Ohshima

Department of Biology, Faculty of Science, Kyushu University, Hakozaki, Higashi-ku, Fukuoka 812, Japan

The small nuclear RNAs (snRNAs) are a class of metabolically stable small RNAs present in the nuclei of eukaryotic cells¹. In mammalian cells, there are six major molecular species (U1 to U6 snRNA), which are complexed with proteins, forming small nuclear ribonucleoprotein particles, snRNPs. Of these, the U1, U2, U4, U5, U6 snRNPs are thought to participate in pre-mRNA splicing as part of the spliceosome (for review, see ref. 2). Here, we describe the characterization of the gene coding for the *Schizosaccharomyces pombe* U6 snRNA. Unexpectedly, the *Schiz. pombe* U6 RNA gene was found to contain an intron-like sequence of 50 base pairs. Northern blot analysis and RNA sequencing revealed that this intron-like sequence is precisely removed from the transcript. The mature U6 RNA of *Schiz. pombe* has 77% sequence homology with the mammalian U6 RNA. In *Schiz. pombe*, it is possible that U6 RNA is not only involved in pre-mRNA splicing, but is also a splicing substrate. This is the first report of an intron in a snRNA gene.

U6 RNA is one of the six abundant, capped snRNAs present in eukaryotic cells³⁻⁹. U6 and U4 RNAs are base-paired and are present in the same snRNP^{10,11}. The U1, U2, U5, U4/U6 snRNPs assemble onto the pre-mRNA substrate to generate the active spliceosome². U4 RNA is thought to be released from the assembled spliceosome during the splicing reaction¹². The U1 snRNP interacts with the 5' splice site, the U2 snRNP with the branch point, and the U5 snRNP may interact with the 3' splice site. The exact function(s) of the U4/U6 snRNP in pre-mRNA splicing, however, is not known.

To elucidate the function(s) of U6 RNA, we isolated the gene coding for U6 RNA of the fission yeast *Schiz. pombe*, which is as convenient for genetic analysis and gene manipulation as the budding yeast *Saccharomyces cerevisiae*. In addition, it has been proposed that the splicing machinery of *Schiz. pombe* is closer than that of *Sacc. cerevisiae* to the machinery of higher eukaryotes¹³.

The nucleotide sequence of U6 RNA is highly conserved among species¹⁴. We synthesized the 30-mer oligonucleotide complementary to the most conserved region of U6 RNA (nucleotides 40-69 of rat U6 RNA⁵) and used this as a probe, initially for Southern and northern blot analysis of the total

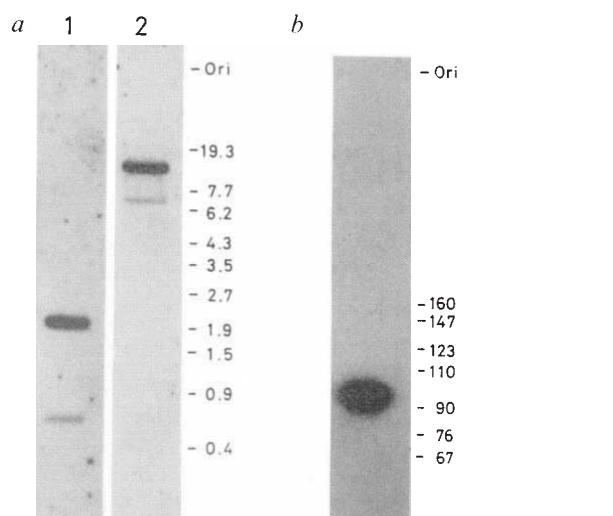


Fig. 1 *a*, Southern blot analysis of total DNA. Total *Schiz. pombe* DNA (10 μ g) was digested with *Hind*III (lane 1) or *Pst*I (lane 2), electrophoresed on a 1% agarose gel, transferred to a Nylon membrane (Biodyne) and hybridized with the kinase-labelled oligonucleotide. Positions of size markers (λ *Sty*I digest) are on the right in kb. *b*, Northern blot analysis. *Schiz. pombe* total RNA (30 μ g) was electrophoresed on a 5% polyacrylamide/7M urea gel, transferred to a Nylon membrane (Biodyne) by capillary action, and probed. Markers were kinase-labelled *Hpa*II fragments of pBR322 DNA.

Methods. Hybridization was at 42 $^{\circ}$ C in a solution containing 6 $\times$ SSC, 50 mM sodium phosphate, 5 $\times$ Denhardt's solution, 0.1% SDS, 2 mM EDTA and 20 μ g ml $^{-1}$ tRNA. The filter was washed in 3 M tetramethylammonium chloride, 2 mM EDTA, 50 mM Tris-HCl (pH 8.0) at room temperature for 30 min and then at 50 $^{\circ}$ C for 20 min. Autoradiography was carried out at -70 $^{\circ}$ C with an intensifying screen.

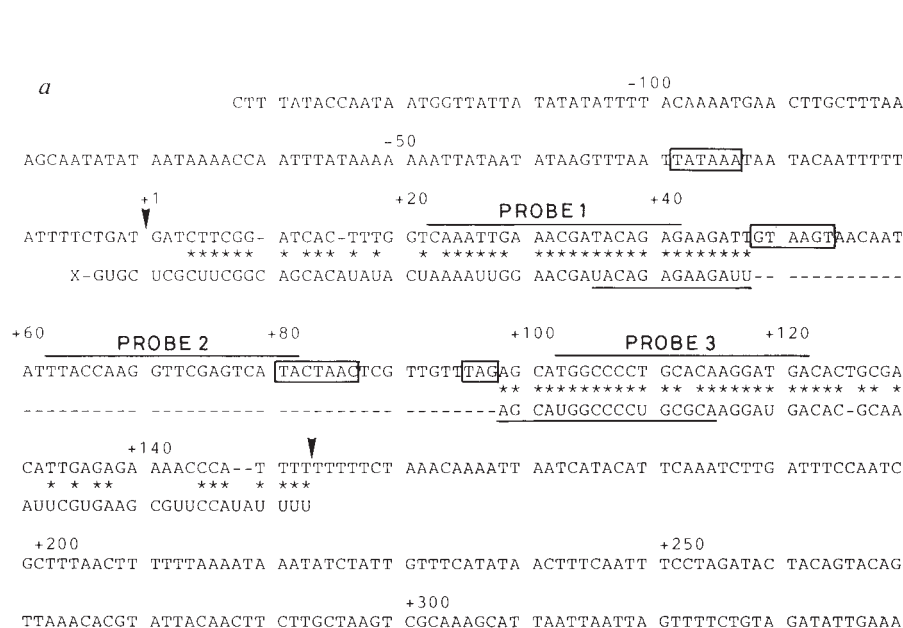


Fig. 2 *a*, Nucleotide sequence of the gene encoding the U6 RNA from *Schiz. pombe*. The upper line shows the sequence of the non-coding DNA strand. The lower line shows the sequence of the rat U6 RNA. Asterisks indicate identical nucleotides and dashes denote gaps introduced to optimize sequence homology. The 5' and 3' ends of the coding region were determined as described in the text and indicated by the vertical arrow heads. A presumptive 'TATA box' and the sequences similar to the consensus sequence of donor, acceptor and branch site are boxed. Numbers indicate distance in nucleotides from the predicted transcription initiation site. Overlines represent the sequences complementary to the probes used for northern blot analysis. Underlines indicate U6 RNA sequence (+40 to +69) complementary to the hybridization probe used for cloning. *b*, Mapping of the 5' end of the coding region. The oligonucleotide complementary to the region +21 to +40 (probe 1) was used for primer extension reactions with total *Schiz. pombe* RNA (lane 5). The products of sequencing reactions using the same primer were run on lanes 1-4. The DNA sequence is written to the left of the panel. Primer extension reactions were performed as described²².

Schiz. pombe DNA and RNA. Autoradiography of a Southern blot showed a single major hybridizing band of 2.0 kilobases (kb) and 9.3 kb and a minor band of 0.8 kb and 6.8 kb in the *Hind*III and *Pst*I digests, respectively (Fig. 1*a*). Similarly, a single major band of 100 nucleotides, close in size to that of mammalian U6 RNA, was detected by northern blot analysis (Fig. 1*b*). These results indicate that a gene which has strong homology with the probe is present at a single copy in the *Schiz. pombe* genome. We therefore constructed a genomic library derived from 1.5-2.5 kb *Hind*III fragments of *Schiz. pombe* DNA, and screened the library with the oligonucleotide probe. Positive clones were analysed by restriction mapping and nucleotide sequencing. All the positive clones analysed con-

tained the same 2.0 kb fragment. Figure 2*a* shows the nucleotide sequence of the region that hybridized with the probe. A single U6 RNA gene was found in the 2.0 kb fragment. The 5' end of the coding region was determined by primer extension (Fig. 2*b*) and RNA sequencing using the dideoxy method (data not shown). The 3' end was tentatively assigned to be at +149 by a comparison with rat U6 RNA sequence. Although all U6 RNA genes cloned to date have a TATA box in the 5'-flanking region, several lines of evidence indicate that U6 RNA is transcribed by RNA polymerase III^{3,6,15,16}. The putative 3' end of the *Schiz. pombe* U6 RNA gene falls within a (T)-residue tract characteristic of transcription termination sites of RNA polymerase III.

Comparison of the nucleotide sequence of the coding region with that of rat U6 RNA revealed an unexpected structure of the *Schiz. pombe* U6 gene. Sequence identity with rat U6 RNA is abruptly interrupted at +47, and is resumed at +97. Interestingly, the 5' and 3' ends of the interruption are GTAAGT and TAG, respectively. These are identical to the consensus sequences at the ends of introns of all eukaryotic genes¹⁷. In addition, the sequence 5'-TACTAAC 3', which is located 12 base pairs (bp) upstream of the 3' boundary, matches the potential branch site sequence 5'-CTRAY-3' present 6–18 residues upstream of the 3' splice site of *Schiz. pombe* genes¹⁸. More remarkably, this sequence is identical to the TACTAAC box, the strictly conserved branch-site sequence found in *Sacc. cerevisiae*^{19,20}. The distance between GT and AG is similar to the length of many *Schiz. pombe* introns. All these structural features suggest that the region +47 to +96 is an intron.

To investigate whether the intron-like sequence is actually an intron, we performed several experiments. First, northern blot analysis was carried out using the oligonucleotides complementary to three parts of the gene as probes (see Fig. 2a). As shown in Fig. 3a, probes 1 and 3 hybridized to RNA of the same size (about 100 nucleotides). In contrast, no hybridizing band was detected with probe 2, which is complementary to the intron-like region. These results suggest that the middle part of the transcript has been removed. Second, the partial nucleotide sequence of the *Schiz. pombe* U6 RNA was determined by the dideoxy method to localize the precise splice site (Fig. 3b). The nucleotide sequence of the mature U6 RNA indicates that the region +47 to +96 is indeed spliced out from the RNA. The U6 RNA sequence determined (68 nucleotides from the 5' end) was completely co-linear with the gene sequence, except for the region +47 to +96. Furthermore, the size of the mature product (99 nucleotides) predicted from the gene sequence agrees well with the result of northern blot analysis. Third, we used Southern blot analysis with 12 restriction enzymes and probe 1 to exclude the possibility that there is an intron-less U6 gene. A single hybridizing band was detected in all cases, confirming that the *Schiz. pombe* U6 RNA gene is unique in the genome (data not shown). The minor bands (0.8 kb and 6.8 kb in *Hind*III and *Pst*I digests, respectively) were not detected with probe 1 even after a long exposure. Therefore, the minor band observed with the 30-mer probe does not seem to represent a U6 gene. Fourth, accumulation of a putative U6 RNA precursor of about 150 nucleotides was observed in *Schiz. pombe* which had been incubated at 43 °C for 30 min before RNA extraction (unpublished results). On the basis of these results we conclude that the *Schiz. pombe* U6 RNA gene has an intron.

The nucleotide sequence of U6 RNAs are highly conserved among species (Fig. 4). The mature *Schiz. pombe* U6 RNA has a 77% sequence identity with mammalian U6 RNA. The intron

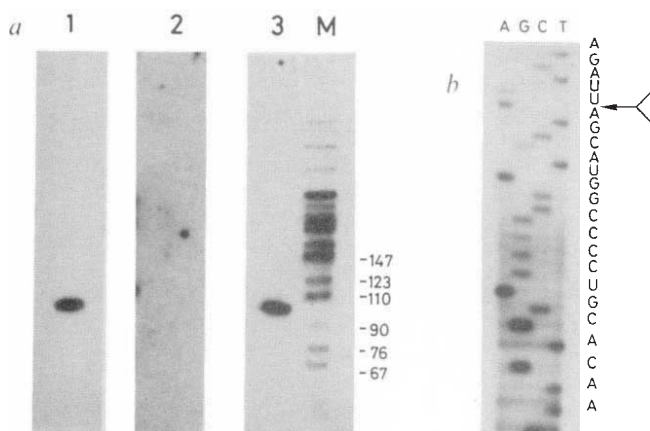


Fig. 3 *a*, Northern blot analysis. *Schiz. pombe* total RNA (30 µg) was electrophoresed on a 5% polyacrylamide/7 M urea gel, transferred to a Nylon membrane (Biodyne), and probed separately with the probe 1 (lane 1), probe 2 (lane 2) and probe 3 (lane 3). Hybridization conditions were the same as described in the Fig. 1 legend. Markers were kinase-labelled *Hpa*II fragments of pBR322 DNA (lane M). *b*, Sequence near the splice junction of the *Schiz. pombe* U6 RNA. The oligonucleotide complementary to the region +121 to +140 was kinase-labelled, and annealed to 30 µg of total *Schiz. pombe* RNA. The sample was split into four aliquots and used to direct cDNA synthesis²² in the presence of ddATP (lane A), ddGTP (lane G), ddCTP (lane C), or ddTTP (lane T). Samples were run on a 8% polyacrylamide/8.3 M urea gel. The RNA sequence predicted by the cDNA sequence is written on the right side of the panel. The arrow shows the position where the intron sequence was located.

is present within the most conserved region of U6 RNA, the region thought to be important for U4:U6 interaction¹⁰.

To our knowledge, this is the first example of an snRNA gene containing an intron. Although various snRNA genes have been cloned, none were found to have an intron. As for U6 RNA genes, introns have never been detected in genes from mammals^{3,21}, *Drosophila melanogaster*⁹, *Xenopus tropicalis*⁶ and the budding yeast *Sacc. cerevisiae*⁷. If *Schiz. pombe* U6 RNA is indeed a RNA polymerase III transcript, as it is in other organisms, the *Schiz. pombe* U6 RNA gene is the first gene to be transcribed by RNA polymerase III to have a typical mRNA intron. In the *Schiz. pombe* U6 RNA gene, the region corresponding to a 'box A' intragenic control sequence (5'-AGATTAG-CATGG-3') of RNA polymerase III is disrupted by the intron

Fig. 4 Comparison of primary structures of U6 RNAs. The nucleotide sequence of the mature *Schiz. pombe* U6 RNA is aligned with that of U6 RNA from the rat⁵, *Drosophila melanogaster*⁹, *Xenopus tropicalis*⁶, *Sacc. cerevisiae*⁷, and broad bean (*Vicia faba* L.)⁸. Colons indicate the nucleotides identical to those of rat U6 RNA. Dashes represent gaps introduced to improve the alignment. The vertical arrowhead indicates the position of the intron present in the *Schiz. pombe* U6 RNA gene.

<i>S. pombe</i>	GAU-----CUUCGG-AUCAC-UUUGGUCAAUUGAAACGAUACAGAGAAGAUUA
rat	GUGCUUG-----CUUCGGCAGCACAUAUACUAAAAUUGGAACGAUACAGAGAAGAUUA
<i>Xenopus</i>	GUGCUUG-----CUUCGGCAGCACAUAUACUAAAAUUGGAACGAUACAGAGAAGAUUA
<i>Sacc. cerevisiae</i>	GUUCGCGAAGUAACCCUUCGU--GGACAUUUGGUCAAUUGAAACAAUACAGAGAUGAUCA
bean	CUUCGG--GGACAUCGGAUAAAAUUGGAACGACACAGAGAAGAUUA
fly	GUUCUUG-----CUUCGGCAGAACAUUACUAAAAUUGGAACGAUACAGAGAAGAUUA
<i>S. pombe</i>	GCAUGGCCCCUGCACAAGGAUGACACUGC-GACAUUGA-GAGAAAA--CCCA--UUUU
rat	GCAUGGCCCCUGCGCAAGGAUGACACG-----CAAAUUCGUGAAGCGUUCUAAUUUU
<i>Xenopus</i>	GCAUGGCCCCUGCGCAAGGAUGACACG-----CAAAUUCGUGAAGCGUUCUAAUUUU
<i>Sacc. cerevisiae</i>	GCAGUCCCCUGCAUAGGAUGAACCGUUUUACAAA--GAGAUUUUUUUG--UUUU
bean	GCAUGGCCCCUGCGCAAGGAUGACACG--CACAAA-UCGAGAAUUGGUCCAAUUUU
fly	GCAUGGCCCCUGCGCAAGGAUGACACG-----CAAAUUCGUGAAGCGUUCUAAUUUU

sequence. This sequence, however, is dispensable for efficient and accurate transcription of the vertebrate U6 RNA gene¹⁵. In addition, there is no consensus polymerase III intragenic sequence in the budding yeast U6 RNA gene, which is transcribed by yeast polymerase III⁷.

It is surprising that the U6 RNA gene of fission yeast, with such a small genome size, has an intron. We do not know whether the intron of the U6 RNA gene has been removed in other organisms, or whether an intron sequence has jumped into the U6 RNA gene of *Schiz. pombe* during evolution. Oligonucleotide probe 2, which is complementary to the intron region gave a single major band corresponding to the cloned gene in Southern blots of the total *Schiz. pombe* DNA (data not shown). This would indicate that the U6 intron sequence is unique in the *Schiz. pombe* genome. It is, therefore, unlikely that the U6 intron is a kind of transposon.

It is possible that the *Schiz. pombe* U6 RNA is involved in RNA splicing as an RNA component of the spliceosome although U6 RNA itself is a substrate for RNA splicing. Probe 2 did not detect a band corresponding to the U6 RNA precursor on a northern blot, suggesting that the precursor is spliced very rapidly. If U6 RNA is a component of the splicing machinery, it must be very efficiently spliced, and thus must have the most efficient splicing signals. It is noteworthy that the intron of the *Schiz. pombe* U6 RNA gene has TACTAAC sequence, identical to the branch-site sequence of *Sacc. cerevisiae*.

The budding yeast gene for U6 RNA (SNR6) is reported to

be essential for cell viability⁷. Gene disruption experiments are in progress to investigate whether the U6 RNA gene is essential for *Schiz. pombe*.

We thank Dr M. Yanagida for providing *Schiz. pombe* strains and M. Ohara for comments. This research was supported by grants from the Ministry of Education, Science and Culture of Japan.

Received 22 August; accepted 9 November 1988.

1. Busch, H., Reddy, R., Rothblum, L. & Choi, Y. C. *A. Rev. Biochem.* **51**, 617-654 (1982).
2. Maniatis, T. & Reed, R. *Nature* **325**, 673-678 (1987).
3. Kunkel, G. R., Maser, R. L., Calvet, J. P. & Pederson, T. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8575-8579 (1986).
4. Harada, F., Kato, N. & Nishimura, S. *Biochem. biophys. Res. Commun.* **95**, 1332-1340 (1980).
5. Epstein, P., Reddy, R., Henning, D. & Busch, H. *J. biol. Chem.* **255**, 8901-8906 (1980).
6. Krol, A., Carbon, P., Ebel, J.-P. & Appel, B. *Nucleic Acids Res.* **15**, 2463-2478 (1987).
7. Brown, D. A. & Guthrie, C. *Nature* **334**, 213-218 (1988).
8. Kiss, T., Antal, M. & Solymosy, F. *Nucleic Acids Res.* **15**, 543-560 (1987).
9. Das, G., Henning, D. & Reddy, R. *J. biol. Chem.* **262**, 1187-1193 (1987).
10. Rinke, J., Appel, B., Digweed, M. & Lüthmann, R. *J. molec. Biol.* **185**, 721-731 (1985).
11. Hashimoto, C. & Steitz, J. A. *Nucleic Acids Res.* **12**, 3283-3293 (1984).
12. Lamond, A. I., Konarska, M. M., Grabowski, P. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 411-415 (1988).
13. Käufer, N. F., Simanis, V. & Nurse, P. *Nature* **318**, 78-80 (1985).
14. Reddy, R. *Nucleic Acids Res.* **16**, r71-r85 (1988).
15. Das, G., Henning, D., Wright, D. & Reddy, R. *EMBO J.* **7**, 503-512 (1988).
16. Reddy, R., Henning, D., Das, G., Harless, M. & Wright, D. *J. biol. Chem.* **262**, 75-81 (1987).
17. Mount, S. M. *Nucleic Acids Res.* **10**, 459-472 (1982).
18. Mertins, P. & Gallwitz, D. *EMBO J.* **6**, 1757-1763 (1987).
19. Langford, C. J., Klinz, F.-J., Donath, C. & Gallwitz, D. *Cell* **36**, 645-653 (1984).
20. Newman, A. J., Lin, R.-J., Cheng, S.-C. & Abelson, J. *Cell* **42**, 335-344 (1985).
21. Ohshima, Y., Okada, N., Tani, T., Itoh, Y. & Itoh, M. *Nucleic Acids Res.* **9**, 5145-5158 (1981).
22. Mayeda, A. & Ohshima, Y. *Molec. cell. Biol.* **8**, 4484-4491 (1988).

Structural differences between a *ras* oncogene protein and the normal protein

Liang Tong, Abraham M. de Vos, Michael V. Milburn, Jarmila Jancarik, Shigeru Noguchi*, Susumu Nishimura*, Kazunobu Miura†, Eiko Ohtsuka† & Sung-Hou Kim‡

Department of Chemistry and Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720, USA

* Biology Division, National Cancer Center Research Institute, Tokyo, Japan

† Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo, Japan

One of the most commonly found transforming *ras* oncogenes in human tumours has a valine codon replacing the glycine codon at position 12 of the normal c-Ha-*ras* gene¹⁻⁴. To understand the structural reasons behind cell transformation arising from this single amino acid substitution, we have determined the crystal structure of the GDP-bound form of the mutant protein, p21(Val-12), encoded by this oncogene. We report here the overall structure of p21(Val-12) at 2.2 Å resolution and compare it with the structure of the normal c-Ha-*ras* protein⁵. One of the major differences is that the loop of the transforming *ras* protein that binds the β -phosphate of the guanine nucleotide is enlarged. Such a change in the 'catalytic site' conformation could explain the reduced GTPase activity of the mutant⁶⁻⁸, which keeps the protein in the GTP bound 'signal on' state for a prolonged period of time, ultimately causing cell transformation.

The cloning and expression⁹ of the synthetic gene coding for human c-Ha-*ras* p21 proteins and the crystallization conditions¹⁰ have been described. Briefly, the synthetic gene (using *Escherichia coli* preferred codons) for amino-acid residues 1-171 was cloned and expressed in *E. coli*. We have deleted the codons corresponding to the carboxy-terminal 18 residues because there

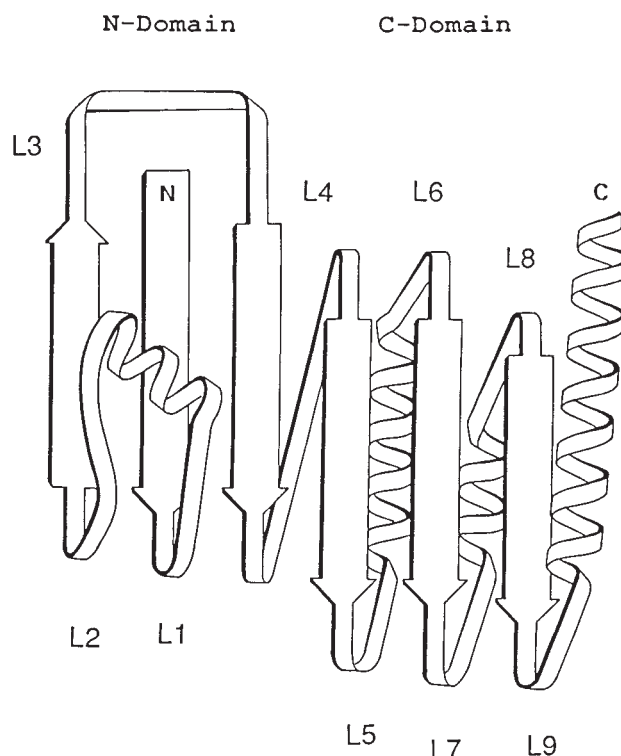
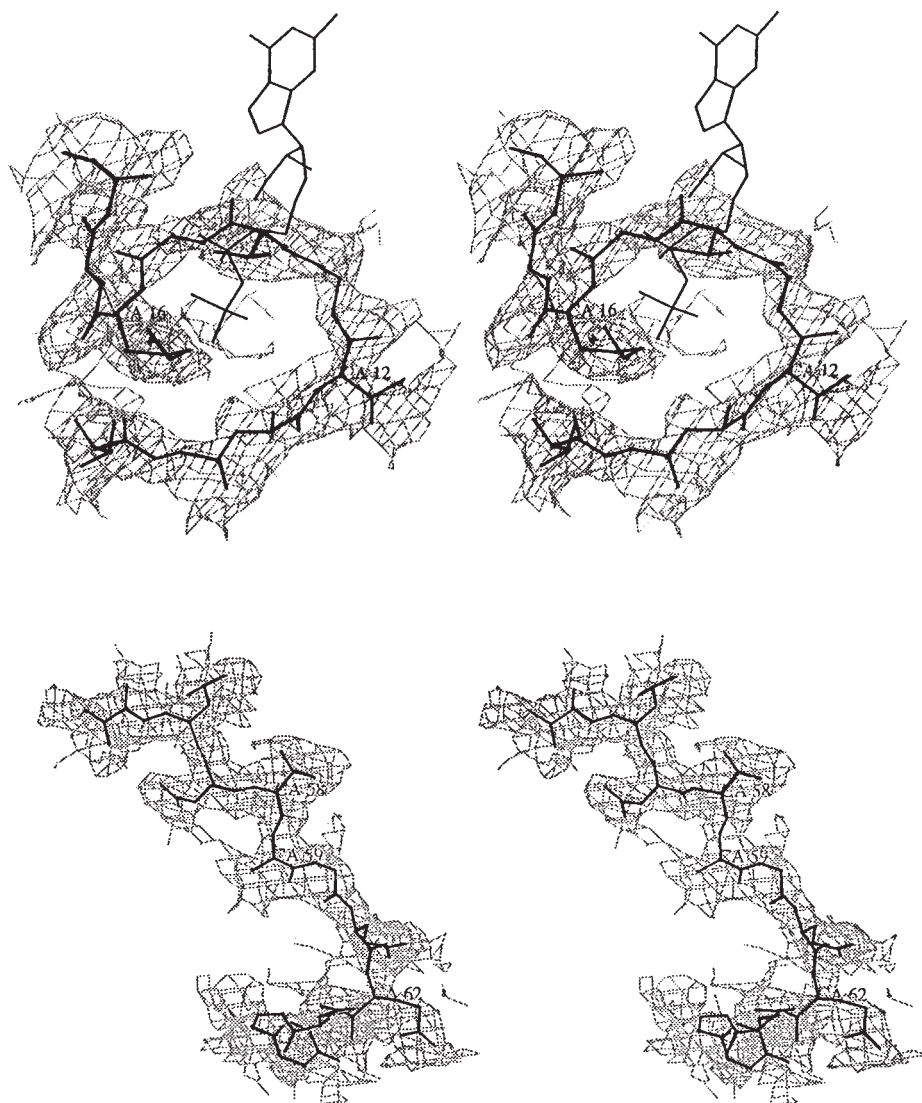


Fig. 1 Topological structure of human c-Ha-*ras* p21 protein. Each β -strand, α -helix and loop is numbered starting from the amino-terminus. The N-domain, which covers three β -strands and one α -helix, is barely hydrogen-bonded to the C-domain. The former has much higher thermal motion than the latter, suggesting that the flexibility of the N-domain is greater. The catalytic site for GTP hydrolysis is localized in the N-domain (phosphate binding domain) and the recognition for guanine base is in the C-domain (guanine recognition domain).

‡ To whom correspondence should be addressed.

Fig. 2 Electron density maps for p21 (Val-12) around loop L1 and part of loop L4 with model fitting. These portions of electron density have been calculated with phases obtained from all atoms in the molecule, omitting those belonging to residues 9–18 and 55–64. The conformation of both regions in this mutant is different from that of normal p21.



is reasonable evidence that this region is flexible^{11,12}, and this flexibility could be detrimental to obtaining good quality crystals. Only the soluble fraction of the protein was purified for crystallization. The GTPase activity of p21(Val-12, 1–171) is less than 10% of that of the normal protein. High-quality single crystals were obtained from solutions containing 10 mg ml⁻¹ protein, 0.1 M CaCl₂, 0.075 M HEPES buffer at pH 7.5, 0.5 mM EDTA, 0.5 mM dithiothreitol and 0.005% *n*-octyl glucoside, equilibrated to 30% PEG 400. The crystals are isomorphous with those of the normal protein, and have space group *P*6₃22 with cell parameters *a* = *b* = 83.2 Å and *c* = 105.1 Å.

The diffraction data for p21(Val-12, 1–171) were collected on an Enraf Nonius rotation camera installed on the 8-pole wiggler line at the Stanford Synchrotron Radiation Laboratory, Palo Alto, California. The X-ray wavelength used for data collection was 1.08 Å, the crystal-to-film distance was 85 mm, and 2° rotation was used for each exposure. One set of data was collected from one crystal at 4 °C on a total of 32 films (16 film packs). The films were digitized on a drum scanner (P-1000, Optronics), and processed with a program originally written by Rossmann¹³. The reflection data from different films were then merged using the PROTEIN program package¹⁴. A total of 48,073 observations were used in the merging process, giving 10,447 unique reflections to 2.2 Å resolution, with an overall *R*(merge) on intensity of 8%.

The crystallographic refinement with the TNT program package¹⁵ used the model for the normal protein⁷ as a starting point.

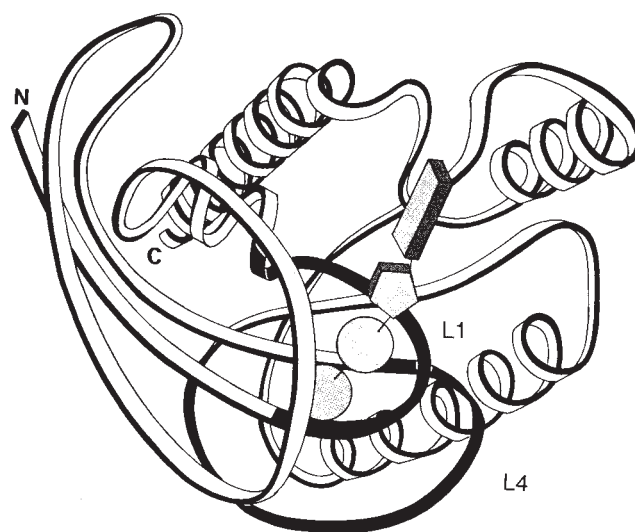


Fig. 3 An artist's drawing, looking down the guanine nucleotide-binding pocket, of the transforming human *c-Ha-ras* oncoprotein p21(Val-12) crystal structure. Cylinders, ribbons and thick black curves represent α -helices, β -strands and loops, respectively. The bound GDP molecule is identified by stippled shapes. The four loops L1, L2, L3 and L4 (Fig. 1) with conformation different from those in the normal protein are labelled.

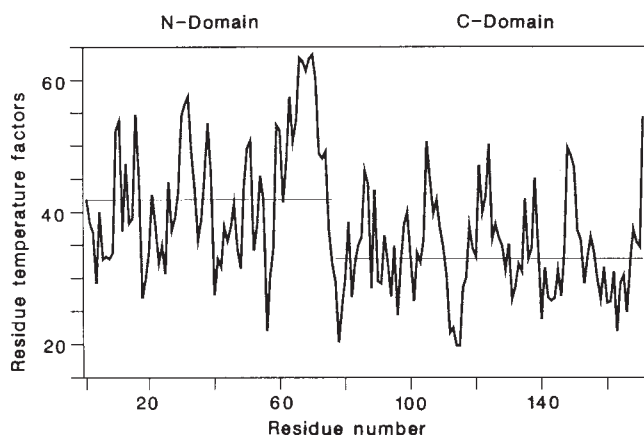


Fig. 4 Average main-chain temperature factors for p21(Val-12). The thin horizontal lines represent the average main-chain temperature factors for the N-terminal (1-75) and C-terminal (76-171) residues. Notice that the temperature factors for residues 60-75 of loop L4 are the highest in the structure.

During the course of refinement, omit maps were calculated by removing residues 9-18 and 55-64, corresponding to loop L1 and part of L4 (Fig. 1), and the model was refitted to the omit maps. Examples of such omit maps are shown in Fig. 2. The current *R* factor for 9,053 reflections between 10 and 2.2 Å resolution (80% of total possible) is 23.6%. A detailed description of the refinement and the refined structure will be reported elsewhere.

The overall structure (Fig. 3) of transforming p21(Val-12) is similar to that of the normal protein⁵. It contains six β -strands, four α -helices, and nine connecting loops (Fig. 1). Both structures seem to consist of two recognizable domains: the amino-terminal domain, containing the first 75 residues (including the first three β -strands and one α -helix), is the phosphate binding domain, and the carboxy-terminal domain, containing the remaining residues (including the last three β -strands and three α -helices), is the guanine recognition domain. There is only a short stretch of hydrogen bonding between β 3 in the N-terminal domain and β 4 in the C-terminal domain (Fig. 1). This separation of domains is also manifested by the distribution of the residue temperature factors in each domain (Fig. 4). The residues in the N-terminal domain have higher temperature factors (an average of 42 Å²), and thus are more mobile than those in the C-terminal domain (an average of 33 Å²). This high mobility of the

N-terminal domain may have a functional significance in that it is in this domain where the catalytic site of GTP hydrolysis, the putative effector region^{16,17}, and the GTPase-activating protein-binding region (residues 30-40)^{18,19} are located.

The differences between the two structures are mainly localized in the loops L1, L2, L3 and L4 in the N-terminal half of the molecule (Figs 1 and 3). The root-mean-square difference in *C α* between p21(Val-12) and the normal protein is 1.26 Å for residues in the N-terminal domain (residues 1-75) and 0.56 Å for residues in the C-terminal domain (residues 76-171). At the current stage of the refinement, the two regions with the largest differences are located in loops L1 and L4. However, the electron density for residues in loop L4 is not as well-defined (as evidenced by their high-temperature factors, Fig. 4). Further refinement is required to assess the significance of the differences in this region.

The clearest and largest differences in structure were found in loop L1, corresponding to residues 9-18, which wraps around the β -phosphate of the bound GDP molecule (Fig. 5). The current refinement results from the normal protein, as well as from transforming p21(Val-12), show that there are several unusual features in this loop (Fig. 5). First, there seem to be no side chains, except that of Lys-16, involved in binding to the phosphate. Second, the backbone amide groups are pointing towards the β -phosphate, making hydrogen bonds to the phosphate oxygens and thus fixing the orientation of the β -phosphate, which is presumably important for GTPase activity. Third, there is a metal ion (probably Mg²⁺ or Ca²⁺) coordinated to the oxygen atoms of the β -phosphate.

The simplest description of the structural differences in loop L1 is that the size of this loop in p21(Val-12) is much larger than in the normal protein (Fig. 5), resulting in the loss of two hydrogen bonds (from the backbone NH groups of residues 12 and 13) to the β -phosphate. We have suggested³ that this loop would have straddled the phosphodiester bond between the β - and γ -phosphates of GTP, and therefore is the prime candidate to be the catalytic site for GTP hydrolysis in the normal p21 protein. The loss of two hydrogen bonds may alter the orientation of the β -phosphate, either when presented to the attacking group, or as a leaving group after the γ -phosphate is attacked, thereby changing the GTP hydrolysis rate. The position of the bound metal ion is such that it is a reasonable candidate for providing the attacking group (probably one of its bound water molecules) to the β -phosphate, perhaps for the on-line displacement of the γ -phosphate. Alternatively, the cation may be participating in fixing the β -phosphate in collaboration with backbone NH groups, with an as-yet-unidentified water molecule attacking the β -phosphate. Another mechanism, similar to that proposed for elongation factor Tu (ref. 20), is

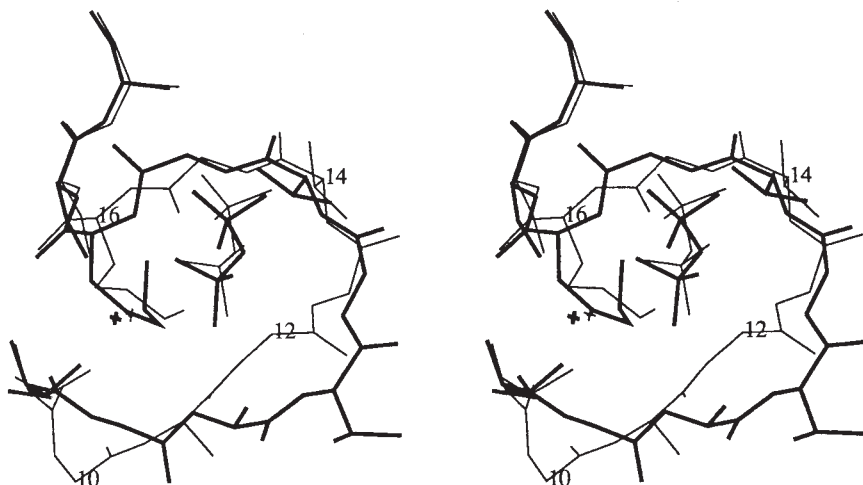


Fig. 5 A stereo view showing the conformation of Loop L1. The thick line is the conformation of loop L1 in p21(Val-12) and the thin line represents that of normal p21 protein. The small crosses represent the positions of the metal ions, and two phosphate groups are shown. Notice that the loop conformation around the β -phosphate in normal p21 protein is much tighter than that of p21 (Val-12).

that the γ -phosphate (which is not present in our structure) is attacked, possibly by a water molecule. This could explain the autophosphorylation activity of the viral *ras* proteins²¹, because in that case the hydroxyl group of Thr-59 of viral p21, located near the presumed γ -phosphate site, could be the attacking group.

In summary, the crystal structure of transforming c-Ha-ras p21(Val-12, 1-171) with a bound GDP molecule provides a plausible explanation for the loss of GTPase activity of the protein, thus keeping the protein in the GTP-bound state, which is thought to signal continued cell proliferation without regulation. A complete understanding will have to await the structure determination of GTP bound c-Ha-ras p21 and of different activated mutants. Efforts to obtain p21 crystals with bound, non-hydrolysable GTP analogue and to determine the structure of the activated mutant p21(Leu-61, 1-171) are in progress. The coordinates will be deposited in the Brookhaven Protein Data Bank.

This work was supported by grants from the National Institute of Health, the Department of Energy, the Japanese Ministry of Health and Welfare, and a gift from Merck, Sharp and Dohme to the University of California at Berkeley. Sh.N. was supported by a fellowship from the Foundation for Promotion of Cancer Research.

Received 7 September; accepted 8 November 1988.

1. Barbacid, M. *A. Rev. Biochem.* **56**, 779-828 (1987).
2. Tabin, C. J. *et al. Nature* **300**, 143-149 (1982).
3. Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149-152 (1982).
4. Taparowsky, E. *et al. Nature* **300**, 762-765 (1982).
5. De Vos, A. M. *et al. Science* **239**, 888-893 (1988).
6. Gibbs, J. B., Sigal, I. S., Poe, M. & Scolnick, E. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5704-5708 (1984).
7. McGrath, J. P., Capon, D. J., Goeddel, D. V. & Levinson, A. D. *Nature* **310**, 644-649 (1984).
8. Sweet, R. W. *et al. Nature* **311**, 273-275 (1984).
9. Miura, K. *et al. Jap. J. Cancer Res.* **77**, 45-51 (1986).
10. Jancarik, J. *et al. J. molec. Biol.* **200**, 205-207 (1988).
11. Shimizu, K. *et al. Nature* **304**, 497-500 (1983).
12. Willumsen, B. M. *et al. EMBO J.* **4**, 2893-2896 (1985).
13. Rossmann, M. G. *J. appl. Crystallogr.* **13**, 225-238 (1979).
14. Steigemann, W. thesis, Technische Univ. München (1974).
15. Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. *Acta crystallogr.* **A43**, 489-501 (1987).
16. Sigal, I. S., Gibbs, J. B., D'Alonzo, J. S. & Scolnick, E. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4725-4729 (1986).
17. Stone, J. C., Vass, W. C., Willumsen, B. M. & Lowy, D. R. *Molec. cell. Biol.* **8**, 3565-3569 (1988).
18. Calés, C., Hancock, J. F., Marshall, C. J. & Hall, A. *Nature* **332**, 548-551 (1988).
19. Adari, M., Lowy, D. R., Willumsen, B. M., Der, C. J. & McCormick, F. *Science* **240**, 518-521 (1988).
20. Eccleston, T. F. & Webb, M. R. *J. biol. Chem.* **257**, 5046-5049 (1982).
21. Shih, T. Y., Stokes, P. E., Smythers, G. W., Dhar, D. & Oroszlan, S. *J. biol. Chem.* **257**, 11767-11786 (1982).

Structure of the membrane-pore-forming fragment of colicin A

M. W. Parker, F. Pattus, A. D. Tucker & D. Tsernoglou

European Molecular Biology Laboratory, Postfach 10.2209, 6900 Heidelberg, FRG

Colicins are antibiotic proteins produced by and active against sensitive *Escherichia coli* and closely related bacteria. They can adsorb to specific receptors located at the external surface of the outer membrane of sensitive cells, and are then translocated to their specific targets within these cells. The largest group of colicins comprises those which can form voltage-dependent channels in membranes, thereby destroying the cell's energy potential¹. Colicin molecules are organized in structural domains, each domain carrying one function associated with the toxin's lethal activity. The pore-forming activity seems to be located at the carboxyl terminus. A thermolytic fragment comprising amino acids 389-592 from colicin A has pore-forming properties very similar to those of the entire molecule. This fragment is soluble in aqueous medium and spontaneously inserts into lipid bilayers². We have determined the

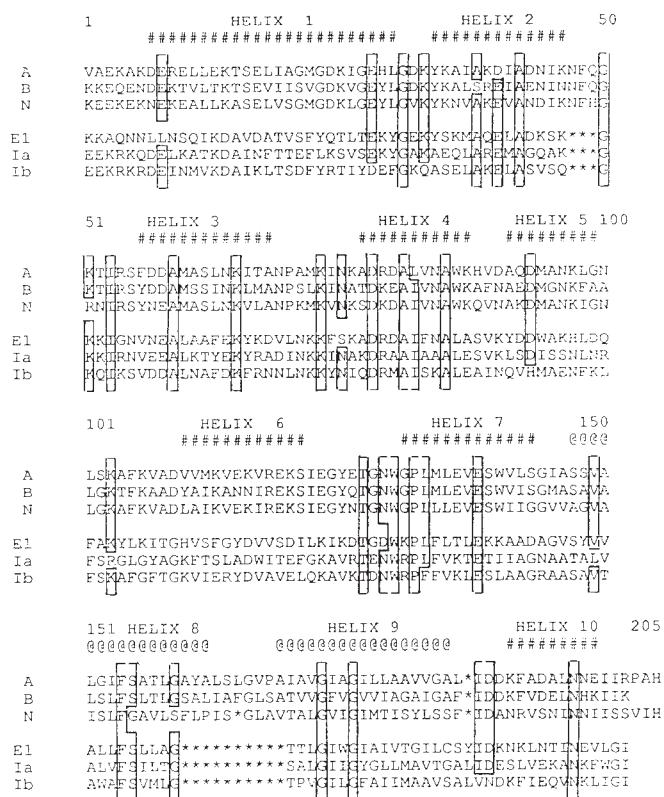


Fig. 1 Aligned sequences of the pore-forming domain of colicins. These sequences can be subdivided into two groups on the basis of sequence identity. Highly conserved residues are boxed. The position of alpha-helices in the sequence are shown. The # symbol refers to surface helices and the @ symbol to buried helices.

structure of the pore-forming fragment of colicin A by X-ray crystallography and refinement at 2.5 Å resolution. The protein consists of ten α -helices organized in a three-layer structure. Two of the helices are completely buried within the structure and form a hydrophobic hairpin loop similar to that proposed for signal sequences which function in translocation. We present a model for insertion of the protein into lipid bilayers the features of which may be applicable in other biological systems involving protein insertion or translocation across membranes.

We have previously reported the crystallization of the thermolytic fragment of colicin A (ref. 3). The crystals belong to space group $P4_32_12$ with cell parameters $a = b = 73.0$ Å, $c = 171.3$ Å. There are two monomers per asymmetric unit. Diffraction data were measured on a multiwire area detector (Nicolet/Xentronics) mounted 20 cm from the crystal on a rotating anode X-ray generator (Elliot GX-18). The data sets were processed using the Genex software package⁴. Over thirty attempts to obtain an isomorphous derivative of the protein crystals using heavy-atom reagents were unsuccessful. Because the wild-type thermolytic fragment of colicin A contains no cysteines, and because mercurial reagents bound to cysteine residues often yield good isomorphous derivatives, we decided to use site-directed mutagenesis to construct cysteine mutants of the proteins (manuscript in preparation). The best heavy-atom derivative was made by soaking in mercury chloride the crystals of a mutant that had a serine replaced by a cysteine (Ser 16 $\rightarrow$ Cys, residue numbering as in Fig. 1) and a repeat decapeptide inserted just before it. The phase set constructed from this derivative was subsequently improved by application of solvent flattening techniques^{5,6}. An atomic interpretation of the electron density map calculated at 3.0 Å resolution was constructed and refined using a combination of molecular dynamics^{7,8} and the distance-

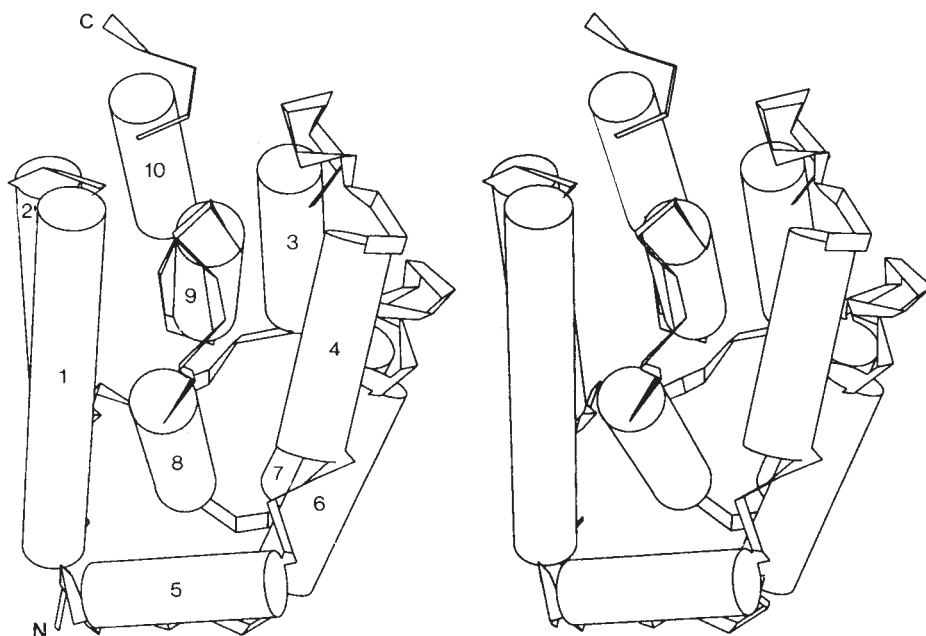


Fig. 2 Stereo representation of the chain fold of the pore-forming fragment of colicin A. Alpha-helices are represented by cylinders and are labelled 1 to 10. This figure was produced by a computer program written by Lesk and Hardman²⁷.

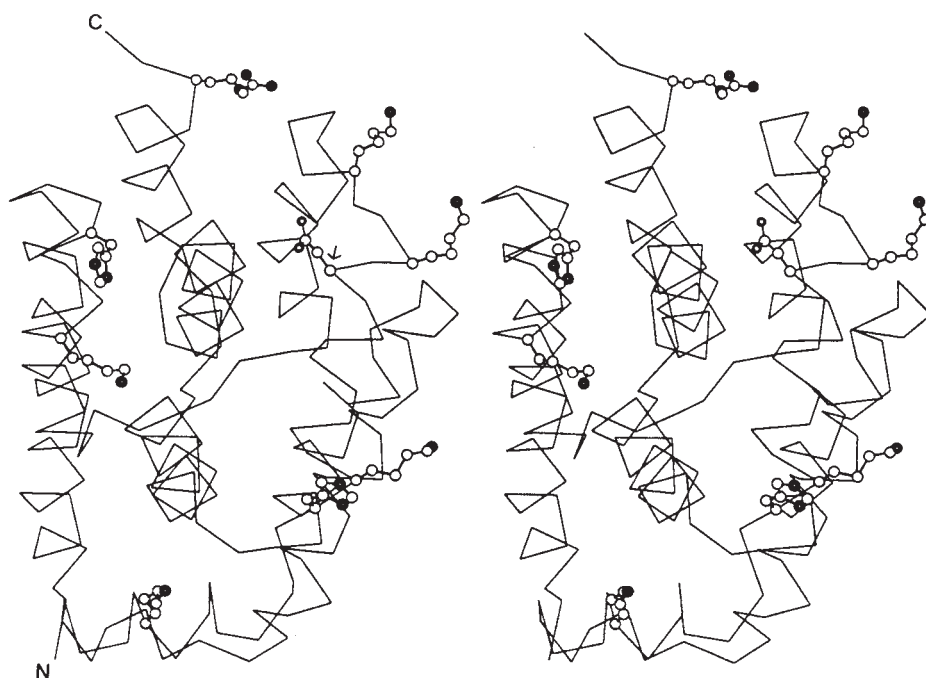


Fig. 3 Stereodiagram of an α -carbon representation of the chain fold showing the array of charged residues on the protein interface thought to interact with membranes. All residues are positively charged with the exception of Asp 78, which is indicated by an arrow. This figure was produced by a computer program written by Lesk and Hardman²⁷.

restrained least-squares procedure of Konnert and Hendrickson⁹. The current crystallographic *R*-factor of the refined model is 24.4% (10,024 reflections $> 3\sigma$, 6.0–2.5 Å, overall temperature factor). The r.m.s. deviation from ideal bond lengths is 0.024 Å. Details of the crystallographic work will be published elsewhere. For ease of discussion the terms 'thermolytic' or 'pore-forming' will be used interchangeably.

The structure of the pore-forming fragment of colicin A is presented in Fig. 2. The molecule can be described as a bundle of ten α -helices which are arranged in three layers. The layer containing the amino terminus of the peptide fragment consists of a pair of antiparallel amphipathic helices (helices 1 and 2). This N-terminal layer is connected by a long loop to another

layer which consists of two pairs of antiparallel amphipathic helices (helices 3 and 4 and helices 6 and 7) with a helix (number 5) connecting the two helix pairs. The middle layer consists of three helices (helices 8, 9 and 10), the first two of which are completely buried and consist of hydrophobic residues only. The peptide fragment crystallizes as a dimer in the asymmetric unit of the unit cell. It should be noted that the protein normally exists as a monomer in solution and the monomeric state is considered to be the physiologically relevant state of the water-soluble form of the protein¹⁰. The dimer interface consists of contacts from the C-terminal end of helices 1, 3, 8 and 10, and from the N-terminal end of helices 2, 4 and 9, as well as contacts from residues in loops connecting these helices. Figure 2 has

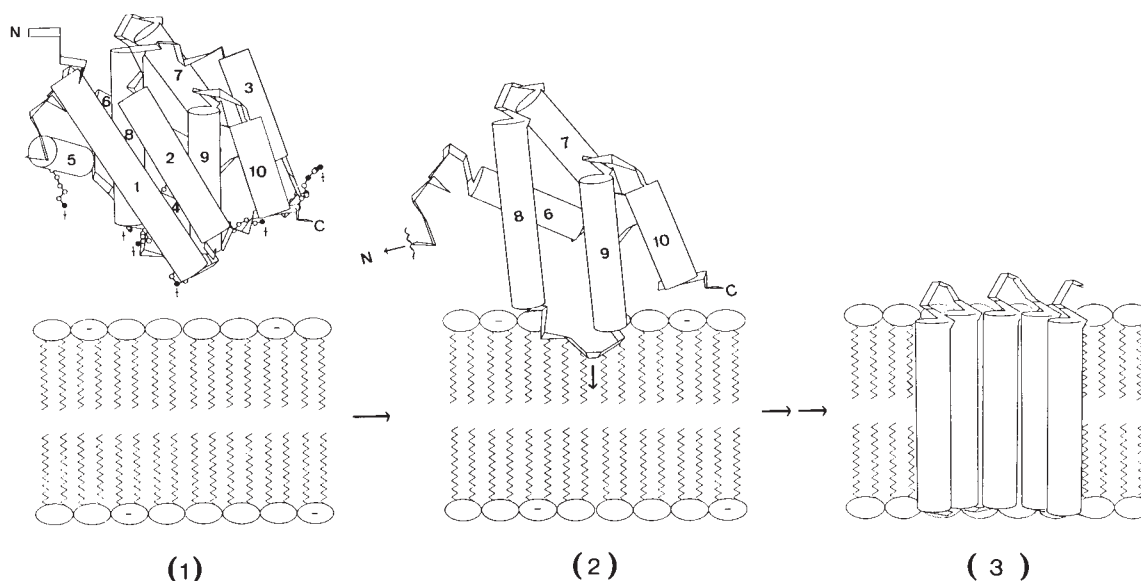


Fig. 4 A model for protein insertion into membranes suggested by the structure of the pore-forming fragment of colicin A. (1) Initial interaction with the membrane occurs as a result of the long-range interaction of an electrostatic potential field and shorter-range interactions such as salt bridges. The first step orientates the hydrophobic helical hairpin (helices 8 and 9) so that its axis is perpendicular to the membrane surface. (2) The hairpin spontaneously penetrates into the membrane, and in doing so initiates a conformational change in the protein. (3) It is envisaged that further steps involve insertion of other helical hairpins which may include oligomerization of the protein molecule. These final steps would lead to the formation of a voltage-gated channel.

been drawn so that this interface appears in the foreground plane of the diagram. The pore-forming fragment of colicin A is the second example of a three-layer α -helical fold first defined by citrate synthase¹¹. There is no resemblance between the two proteins as far as structural details or function are concerned however.

The paradox that a colicin can exist as both a water-soluble protein and as a membrane protein is resolved by the structure of the pore-forming fragment of colicin A. Colicin A and its thermolytic fragment are soluble in water and monomeric over a wide range of *pH* and concentration conditions¹⁰. But the sequences of colicin A and related colicins all contain a hydrophobic stretch of residues between 35 and 49 amino acids long (Fig. 1). This stretch of sequence would be classified as a membrane-spanning sequence on the basis of hydropathy profiles¹². In the three-dimensional structure, this membrane-spanning sequence adopts a helical hairpin motif (helices 8 and 9) which is completely buried within the protein. Thus residues that interact with lipid in the membrane-bound form of the protein are buried within the protein conformation of the water-soluble form.

The mechanism of insertion of a protein into a lipid bilayer can be divided into two steps: binding to the membrane surface and insertion of the protein into the hydrophobic core of the bilayer. The first step is probably electrostatic in nature, whereas the second may be driven by the partitioning of hydrophobic residues into the aliphatic environment of the bilayer. The electrostatic nature of the binding is supported by the following observations. Colicin A and its thermolytic fragment do not bind to neutral bilayers but display a great affinity for negatively charged lipids¹³. Binding can occur without insertion below the lipid phase transition. Penetration of the bilayer is optimized at acidic *pH* (the structure of the thermolytic fragment was determined at *pH* 4.4). The pore properties of colicin A suggest it is asymmetrically orientated with respect to the membrane. Analysis of the distribution of surface side chains of the pore-forming peptide structure shows that they are randomly distributed with respect to various potential interaction parameters such as charge, hydrophobicity and aromaticity. The exception to this finding occurs at the surface defined by the loop that connects

helices 8 and 9 in the structure (Fig. 2). The loop is surrounded by a ring of eight positively charged side chains (residues 25, 29, 73, 76, 87, 88, 97 and 201; see Fig. 3). The only negatively charged residue on this surface is Asp 78 which is located inside the ring of positively charged residues (Fig. 3). It is striking that the interface defined by this loop and surrounding side chains is the same interface used in formation of the non-crystallographic dimer observed in the unit cell of the crystal. Thus this interface appears to have the potential to bind to surfaces found on proteins or lipid bilayers. The mean positive charge of the interface explains the affinity of colicin for negatively charged lipids.

The proposed orientation of binding of the pore-forming domain positions the hydrophobic helical hairpin directly above the membrane surface with helices 8 and 9 sitting perpendicular to the plane of the bilayer. Figure 2 has been drawn such that the proposed interaction surface of the peptide with the bilayer is placed in the foreground plane of the diagram. The positioning of the helical hairpin and the hydrophobic nature of the exposed loop connecting the two helices together with the loss of solvation at the peptide-membrane interface on binding promotes the insertion of the hairpin into the hydrophobic core of the bilayer. This implies that the structure must undergo significant conformational changes because the hairpin constitutes the hydrophobic core of the water soluble form of the protein. There is some evidence that supports a partial unfolding of colicin upon insertion into bilayers. Lowering the dielectric constant of the aqueous solvent with ethanol promotes the partial collapse of the far ultraviolet circular dichroism spectrum of the pore-forming peptide, whereas further addition of ethanol reverses the trend¹². A similar effect is observed when the peptide is added to various concentrations of dimyristoylphosphatidyl glycerol vesicles at acidic *pH* (F. P., unpublished results). There is evidence of a drastic change in the mobility of some side chains on binding to lipid vesicles, as observed by circular dichroism spectroscopy in the near ultraviolet region (F.P., unpublished results). Proteolysis experiments indicate that the C-terminal helix does not get inserted, but is retained on the insertion side of the lipid bilayer¹⁴.

At this stage there are not enough data to predict the protein

topology when inserted into the membrane. But it is noteworthy that the structure of the pore-forming fragment shows that helices are organized in a pairwise fashion. It is tempting to speculate that the insertion occurs by a stepwise insertion of helical hairpins without large changes in secondary structure conformation. The α -helical content of the water-soluble and lipid-bound form of the peptide fragment is identical, as judged by circular dichroism¹². Mutation studies indicate that as few as 130 amino acid residues from the C-terminal end of the peptide are sufficient for pore formation¹⁵. A similar study of the homologous colicin E1 indicates that only the last 88 amino acids are required for pore formation¹⁶. These studies suggest that the insertion of the first helical hairpin (helices 8 and 9) into the membrane provides the energy for insertion of other helical hairpins. After insertion of the first helical hairpin, a large degree of hydrophobic association energy between the other helical hairpins would be lost, making the insertion of further hairpins more energetically favourable. The insertion of the first hairpin would bring the other hairpins closer to the membrane surface and hence make them more responsive to any electrochemical potential effects. Model-building suggests that the pore is probably formed by an oligomer of colicin molecules. An oligomerization process could have an important role in the insertion of other helical hairpins by shielding their polar faces from contact with the hydrophobic core of the membrane and forming the water-filled pore of the voltage-gated channel in doing so. Key features of the model for membrane insertion are presented in Fig. 4.

The three-dimensional structure of the pore-forming fragment of colicin A reveals a number of features relevant to the mechanism of protein insertion into membranes. The last sixty-odd residues at the extreme C-terminus of the protein show many features reminiscent of signal peptides that target proteins to bacterial plasma membranes or to the endoplasmic reticulum of eukaryotic cells. These features include a long stretch of hydrophobic residues with a centrally located glycine (residue 166), followed by a charged C-terminus containing basic residues¹⁷ (Fig. 1). The signal-like peptide adopts a predominantly α -helical structure in colicin A, which is in accord with circular dichroism results for many signal peptides. The hydrophobic part of the signal-like peptide is almost completely buried in the protein, which provides an energetically favourable situation in which a signal peptide can be transferred through a water-soluble medium. It is known that deletions in the regions of the surface helices of the peptide fragment result in precipitation of such mutants in the cytoplasm of the producing cell and thus impede colicin export¹⁵. It follows that the shielding of the hydrophobic region of signal peptides in the water-soluble medium of a cell is an important feature of transported proteins. The helical hairpin structure of the signal-like peptide supports earlier theories that helical hairpin motifs would be a major feature in models of spontaneous insertion of proteins into membranes¹⁸. The structure resolves the paradox of how a water-soluble protein can also be incorporated into membranes. Surfaces required to interact with the hydrophobic core of membranes are buried within the conformation of the water-soluble form of the protein: colicin A is an 'inside-out' membrane protein. This result is consistent with other recent experimental data indicating that proteins require a flexible conformation in translocating across membranes¹⁹⁻²⁴.

The features suggested as being necessary for insertion of colicin A into membranes seem to apply, with some additional requirements, to a wide variety of biological systems. Prime examples include the membrane-attack complex of the complement cascade²⁵, the action of various toxins²⁵ and insertion of membrane precursors²⁶. Current work is directed towards determining the structure of the membrane-bound form of colicin A to understand further the mechanism of insertion and to explain the pore properties of the voltage-gated channel formed.

We thank N. Didat and M.-P. Reck for technical assistance,

Drs D. Baty, M. Chartier and C. Lazdunski for their role in the site-directed mutagenesis work, Dr P. Tucker for help with the X-ray diffraction data collection, and Drs B. Dobberstein, J. Lakey and S. Fuller for critical reading of the manuscript. We also thank C. Barber for assistance in preparation of the manuscript.

Received 22 September; accepted 14 November 1988.

1. Lazdunski, C. *et al. Biochim. biophys. Acta* **947**, 445-464 (1988).
2. Collarini, M., Amblard, G., Lazdunski, C. & Pattus, F. *Eur. biophys. J.* **14**, 147-153 (1987).
3. Tucker, A. D., Pattus, F. & Tsernoglou, D. *J. molec. Biol.* **190**, 133-134 (1986).
4. Howard, A. L. *et al. J. appl. Crystallogr.* **20**, 383-387 (1987).
5. Wang, B. C. in *Meth. Enzym.* **115**, 90-112 (1985).
6. Leslie, A. G. W. *Acta crystallogr.* **A43**, 134-136 (1987).
7. Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458-460 (1987).
8. van Gunsteren, W. F. *Protein Eng.* **2**, 5-13 (1988).
9. Hendrickson, W. A. & Konert, J. H. *Biomolec. Struct. Funct. conform. Evol.* **1**, 43-57 (1981).
10. Cavard, D. *et al. Eur. J. Biochem.* **172**, 507-512 (1988).
11. Remington, S., Wiegand, G. & Huber, R. *J. molec. Biol.* **158**, 111-152 (1982).
12. Pattus, F., Keitz, F., Martinez, C., Provencher, S. W. & Lazdunski, C. *Eur. J. Biochem.* **152**, 681-689 (1985).
13. Pattus, F. *et al. Biochemistry* **22**, 5698-5703 (1983).
14. Cleveland, M. B., Slatin, S., Finkelstein, A. & Levinthal, C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2589-2593 (1977).
15. Baty, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 1152-1156 (1987).
16. Liu, Q. R. *et al. Proteins* **1**, 218-229 (1986).
17. Lunn, C. A., Takahara, M. & Inouye, M. *Meth. Enzym.* **125**, 18-149 (1986).
18. Engelman, D. & Steitz, T. A. *Cell* **23**, 411-422 (1981).
19. Schleyer, M. & Neupert, W. *Cell* **43**, 339-350 (1985).
20. Eilers, M. & Schatz, G. *Nature* **332**, 228-232 (1986).
21. Randall, L. L. & Hardy, S. J. S. *Cell* **46**, 921-928 (1986).
22. Zimmermann, R. & Meyer, D. I. *Trends biochem. Sci.* **11**, 512-515 (1986).
23. Vestweber, D. & Schatz, G. *EMBO J.* **7**, 1147-1151 (1988).
24. Eilers, M., Hwang, S. & Schatz, G. *EMBO J.* **7**, 1139-1145 (1988).
25. Bhakdi, S. & Tranum-Jensen, J. *Rev. Physiol. Biochem. Pharmac.* **107**, 148-106 (1987).
26. Wickner, W. *Biochemistry* **27**, 1081-1086 (1988).
27. Lesk, A. M. & Hardman, K. D. *Science* **216**, 539-540 (1982).

GUIDE TO AUTHORS

GENERAL

Please follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers for whom English is a second language and those who work in other fields. Please write clearly and simply, avoiding unnecessary technical terminology. *Nature's* staff will edit manuscripts to those ends if necessary. Contributors should check their proofs carefully. Because of the competition for space, many of the papers submitted for publication cannot be accepted. For this reason, and because brevity is a great assistance to readers, papers should be as brief as is consistent with intelligibility. Please note that one printed page of *Nature*, without diagrams or other interruptions of the text, has fewer than 1,300 words.

CATEGORIES OF PAPER

Review Articles survey recent developments in a field. Most are commissioned, but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance but must not exceed six pages of *Nature*.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text or more than six display items (figures and tables) and should not occupy more than five pages of *Nature*.

Articles should be accompanied by a heading of 50-80 words written to advertise their contents in general terms, to which editors will pay particular attention. A heading (printed in italic type) is not an abstract and should not usually contain numbers or measurements. The study should be introduced in more detail in the first two or three paragraphs, which should also briefly summarize its results and implications.

Articles may contain a few subheadings of two or three words. The meaning of the text should not depend on the subheadings, whose function is to break up the text and to point to what follows. There should be fewer than 50 references.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should not have more than 1,000 words of text or more than four display items and should not occupy more than two pages of *Nature*. The first paragraph should describe, in not more than 150 words, the origins and chief conclusions of the study. Letters should not have subheadings or more than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned, most are unsolicited. They are normally between one and four pages of *Nature* in length.

News and Views articles are intended to inform non-specialist readers about a recently published advance. Suggestions should be made to the News and Views Editor. Illustrations are welcome. Proposals for meeting reports should be agreed in advance.

Scientific Correspondence is for the discussion of scientific matters, including contributions published in *Nature*. Priority is given to contributions of less than 500 words and five references. Figures are welcome.

Submission. All manuscripts may be submitted to either London or Washington. They should not be addressed to editors by name. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax (VAT).

All contributions submitted for publication in *Nature* should conform with these rules.

Manuscripts should be typed, double-spaced, with a good-quality printer, on one side of the paper only. Four copies are required, each accompanied by lettered artwork. Four copies of half-tones should be included. Reference lists, figure legends and so on should be on separate sheets, all of which should be double-spaced and numbered. Copies of relevant manuscripts in press or submitted for publication elsewhere should be included, clearly marked as such. Revised and resubmitted manuscripts should also be clearly marked as such and labelled with their reference numbers.

Titles should say what the paper is about with the minimum of technical terminology and in fewer than 80 characters. Authors should avoid active verbs, numerical values, abbreviations and punctuation and should include one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name, figure number and, when known, manuscript reference number. Original artwork should be unlettered. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are permitted only if the parts are closely related, either experimentally or logically. Suggestions for cover illustrations, with captions, are welcome. Original artwork (and one copy of the manuscript) will be returned when a manuscript cannot be published.

Colour Artwork. A charge of £500 a page is made for 4-colour figures. Inability to meet these costs will not prevent the publication of essential colour figures if the circumstances are explained.

Figure legends must not exceed 300 words and ideally should be much shorter, and should use telegraphic form wherever possible. The figure should be described first, then, briefly, the method. Reference to a method described elsewhere is preferable to a full description. Methods should not be described in detail in the text.

References should be numbered sequentially as they appear in the text, followed by those in tables and finally those in the figure legends. Reference numbers apply only to papers published or in the press, and a different number should be given to each paper cited. All other forms of reference (including unreferenced abstracts) should be included in the text as personal communication, manuscript in preparation or preprint (with number and institution where appropriate). Text should not be included in the reference list. References should be abbreviated according to the *World List of Scientific Periodicals* (fourth edition, Butterworth, London, 1963-65). The first and last page numbers should be cited. Reference to books should clearly indicate the publisher and the date and place of publication.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided as much as possible and, when used, defined. Editors will shorten words if necessary. In case of doubt, SI units should be used.

Footnotes should not be used except for changed addresses or to identify the corresponding author if different from the first named.

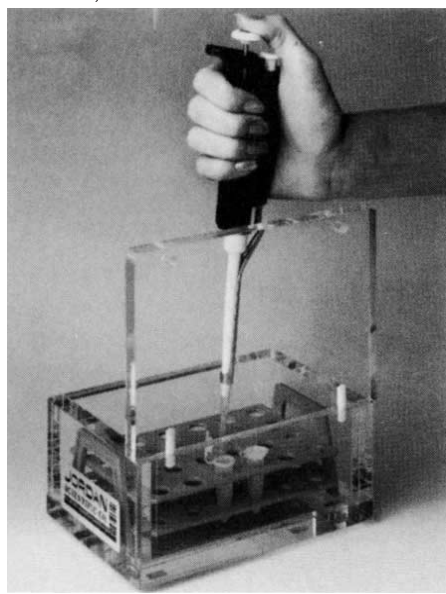
Acknowledgements should be brief. Grant numbers and contribution numbers are not allowed.

New products for a new year

Start the year off with a new addition to the laboratory — this week's selection includes an infrared hotplate, a software model for teratogenicity studies, and a balance that calculates molarity.

BIO MEDICAL Equipment offers what may well be the most inexpensive **electrophoresis apparatus** on the market: the Mac Gel electrophoresis kit (*Reader Service No. 100*). For a total price of \$10 (US), the Mac Gel kit contains two gel-forming plates, a multiwell comb, and platinum electrodes. Bio Medical Equipment's kit prepares standard-sized gels, and the entire apparatus fits into a buffer tank with ample dimensions for large buffer volumes. It is possible to use the electrodes with virtually any controller.

From Jordan Scientific comes a micro-centrifuge **tube rack/beta shield** with a unique two-position lid (*Reader Service No. 101*). With the lid down, the rack is



Storage and safety are the two objectives fulfilled by Jordan Scientific's tube box.

fully enclosed for the safe storage and transportation of tubes containing beta-emitting isotopes. In the upright position, the lid shields the user from radiation during pipetting operations or tube handling. The tube rack is made of one-half-inch-thick acrylic, and can withstand autoclaving. It holds 24 1.5-ml Eppendorf-type tubes in a four-by-six array, a configuration Jordan Scientific says is ideal for performing sequencing reactions. The tube rack costs \$55 (US).

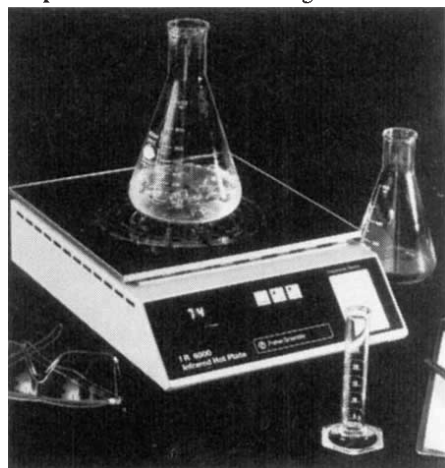
Lab essentials

The newest **spectrophotometer** from Pharmacia LKB is the Novaspec II, an instrument that the company terms "simplicity itself" (*Reader Service No. 102*).

The compact Novaspec II performs absorbance and transmission measurements in only three steps: the required wavelength is set using the keypad, the reference blank is adjusted by one key-stroke after inserting the blank cuvette, and the sample measurement is read off of the digital display. The instrument can make concentration measurements, perform simple kinetics calculations, and measure absorbance ratios and differences. It comes with a bi-directional RS232C interface for connection to a printer for hard-copy results. The Novaspec II's wavelength range is 325–900 nm, monitored by a single detector. It accepts both 10 mm cuvettes and standard-sized test tubes.

Sartorius has just announced the launch of a new series of **analytical balances** that can perform the necessary calculations for preparing molar solutions (*Reader Service No. 103*). The balances come equipped with Sartorius's Lab-Plus software, which has programs for statistical analyses, density determination, and classification and sorting operations. For preparing molar solutions, the operator keys in the molecular weight of the substance to be dissolved, the desired molarity and the target volume — the balance displays the amount of substance to be "filled toward zero". If the operator adds too much, the balance recalculates the target volume to attain the same molarity. The balances can weigh directly in moles, percent by volume or p.p.m. when provided the molecular weight and target volume. The capacity of the balances is 200 g, displayed down to 0.1 mg.

Fisher Scientific has a new **laboratory hotplate** that uses infrared light instead of

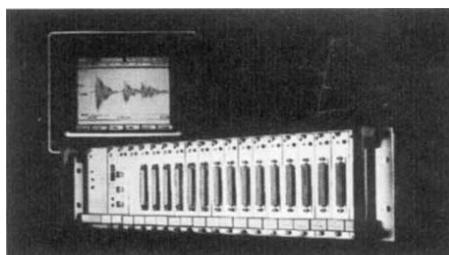


Heating through IR light from Fisher.

an electrical heating unit, to provide a maximum temperature of 615 °C (*Reader Service No. 104*). The company says the hotplate can boil 1,000 ml of water in under 12 minutes in a 1.5-litre Erlenmeyer flask — nearly twice the heating rate of most other hotplates. The \$650 (US) IR-6000 hotplate is continuously adjustable over its power setting range of between 1 per cent and 99 per cent. Fisher Scientific says the hotplate assures greater uniformity of heating than traditional hotplates because it does not depend on the shape of the vessel bottom: the infrared light generated by the unit's molybdenum disilicide coil heats the sample directly. The glass-ceramic surface of the IR-6000 measures 12 × 12 inches, and its stainless steel edge can contain up to 400 ml of liquid in the event of a spill.

Computing

For those who collect large volumes of analogue data from laboratory instru-



Biodata's Microlink data collection system can handle 25,000 samples per second.

ments or transducers with voltage outputs, Biodata offers its Microlink high-speed **data collection system** (*Reader Service No. 105*). Biodata's system collects and streams data directly to the hard disk of an IBM microcomputer or compatible at rates of up to 25,000 samples per second. The data is available immediately for analysis, and the unit can store up to 16 million samples. The data acquisition hardware accepts up to 256 channels in one cabinet and incorporates 12-bit buffered analogue-to-digital conversion. Both voltage and thermocouple inputs can be used with the £4,000 (UK) Microlink system, which includes hardware, software and IBM-compatible computer.

The HDi corporation offers an alternative to animal testing for **teratogen analysis** — a set of advanced structure-activity (SAR) models of teratogenesis coupled with its Topkat software (*Reader Service No. 106*). HDi says the models and the Topkat program can predict the teratogenicity of untested chemicals with an

accuracy of 96 per cent. The models are based on a data base of teratogenesis bioassays that HDi abstracted from the literature. Independent parameters in the SAR equations include electronic charges, substructural keys and topological descriptors. The model is based on four subequations, but this is transparent to the user: Topkat automatically selects the equation appropriate for the chemical to be estimated. HDi says the \$15,000 (US) system correctly classifies carbocyclic aromatics, heterocyclic aromatics, acyclic aliphatics and alicyclics over 95 per cent of the time. Topkat also has modules to predict the probability of carcinogenicity, mutagenicity, skin and eye irritation, and values of acute toxic effects.

An interactive **graphics software package** for exploring, analysing, and visualizing technical data is new from Precision Visuals (*Reader Service No. 107*). PV-WAVE stands for Precision Visuals' Workstation Analysis and Visualization Environment. It allows researchers to navigate through large datasets, reduce and filter the data into manageable subsets, visually distinguish and analyse key features and trends, and translate results into publication-quality charts, graphs, contour maps, surface plots and images. Commands can be entered using the keyboard or be grouped into procedure files; application developers can create custom systems using PV-WAVE's menu development tools. Running on a VAXstation 2000, Precision Visuals says the software can read 2,400 floating-point numbers in 1.6 seconds, or compute a 4,096-point real-to-complex Fast Fourier Transform in 0.15 seconds. PV-WAVE costs \$3,350 for the VAXstation 2000. Sun workstation versions will be released early this year.

A desktop **DNA sequence analysis software package** for use with Macintosh computers has been introduced by International Biotechnologies, Inc. (*Reader Service No. 108*). The \$2,295 (US) IBI Pustell Software for the Macintosh supports over 100 applications, including restriction analysis, homology matrices and protein structural studies. It also provides a global search function that matches DNA-DNA, protein-protein and protein-DNA with designated databases. The software interacts with and converts



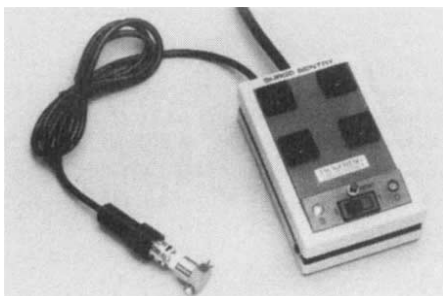
DNA sequence analysis comes to the Macintosh computer with software from IBI.

to every major sequence format, says IBI, including that used in GenBank, Bionet, Staden, Line and UWGCG. A unique "wildcard" feature allows researchers to access data without complete reference information. IBI says researchers can use data from their Macintosh in a VAX or IBM-PC through a user-supplied interface, permitting users to share data with colleagues.

Analytical aids

Spectramass Ltd's BioQuad gas analysis system uses quadrupole mass spectrometry technology to **monitor fermentation processes** (*Reader Service No. 109*). The #15,000-19,000 (UK) system analyses inlet and exhaust gases, and provides data on dissolved gases and volatile substrates. The heart of the system is the Spectramass PC 2000 mass spectrometer, which has been designed with ion source optics especially suited for the analysis of complex gas streams. Spectramass says the analysis of dissolved gases and volatiles is easily accomplished using a membrane inlet system and probe to permit connection to standard fermenter ports. The Bio-Quad system includes vacuum components, remote-controlled valves and inlet lines, and software.

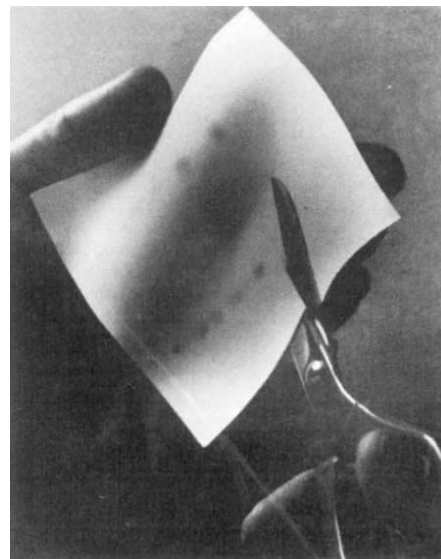
To protect HPLC equipment against damage from reagent backflow, Pickering Laboratories has designed its post-column **eluent flow interlock** (*Reader Service No.*



Prevent post-column eluent backflow with the interlock device from Pickering Labs.

110). The interlock system monitors pressure at the outlet of the eluent pump, and shuts off the power to the post-column system if the eluent pressure drops below 500 p.s.i., stopping reverse reagent flow that can damage both the analytical column and post-column equipment. Pickering Labs' interlock also shuts down the post-column system in the event of power loss, and a surge protection feature guards against electrical damage from voltage spikes. The \$420 (US) interlock comes equipped with four 120-V receptacles, status lights, reset button, circuit breaker, power cord and a pressure switch with a five-foot cord.

Analytichem has introduced a **thin-layer chromatography sheet** that it says is so flexible that spots can be cut or punched



The durable TLC sheets from Analytichem won't crumble during handling.

out for further analysis without resorting to tedious scraping (*Reader Service No. 111*). Analytichem's Empore TLC sheets were developed with the 3M corporation, and are composed of inert PTFE fibres with chemically bonded silica particles. Because the sheets are produced without a binder, Analytichem says they can be used in a wide range of applications with optimum resolution and reproducibility. The company also says the Empore TLC sheets exhibit several times the sample capacity of a standard hard-layer TLC plate, preventing problems with overloading and spot distortion. The 2.5 × 10 cm and 10 × 10 cm sheets are available with a variety of bonded silicas.

Enzymes and reagents

Janssen Life Sciences Products has a new ultra-small **immunogold probe** for increased labelling efficiency and penetration power in the detection of antigens and biotinylated probes using the immunogold silver staining technique (*Reader Service No. 112*). The AuroProbe One reagent contains colloidal gold particles measuring 1 nm, and is designed so that the ratio between probe protein and gold particle is less than one. Because AuroProbe One can be used in a variety of applications, from immunoblotting to electron microscopy, it allows results from different methods to be compared. The signal produced by AuroProbe One can be visualized by silver enhancement using Janssen's IntenSE reagents for microscopy and blotting experiments. Each package of AuroProbe One contains 1 ml AuroProbe One reagent, 10 ml IGSS-quality gelatin and detailed instructions.

Amersham has recently introduced its Multiprime system for **rapid hybridization reactions** (*Reader Service No. 113*). The new kit is based on Amersham's Multiprime random priming labelling kit, and

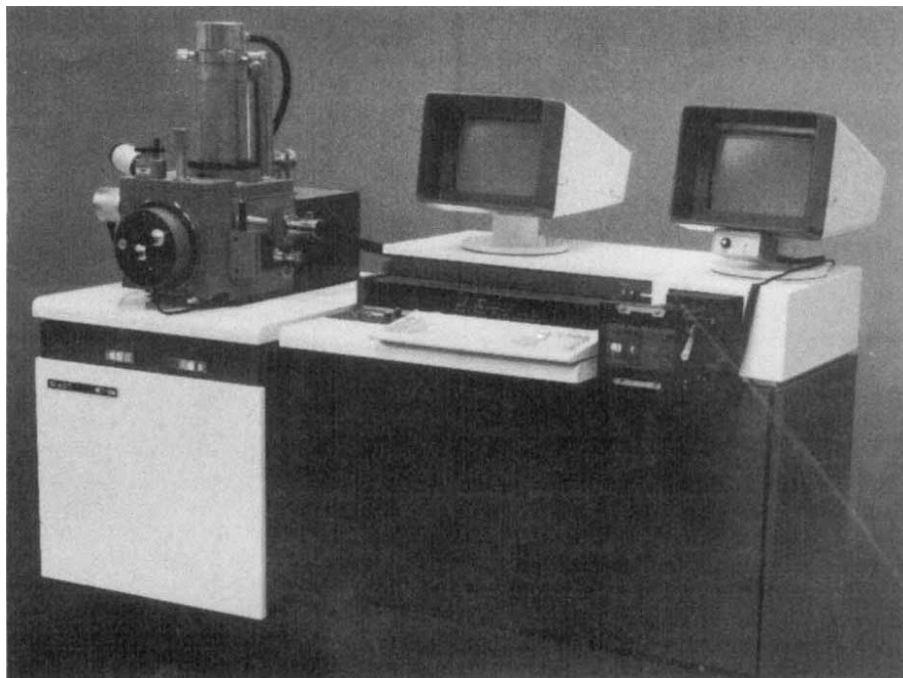
allows the synthesis of high specific-activity probes in 30 minutes, and subsequent hybridization with nucleic acid blots in two hours, says the company. The system consists of a DNA labelling kit and a rapid hybridization buffer that increases the hybridization rate without requiring high probe concentrations. Amersham says when using a probe concentration of between 2 and 8 ng ml⁻¹, the sensitivities obtained after a two-hour hybridization are equivalent or higher than sensitivities obtained in an overnight incubation using conventional buffers, without an increase in background.

GIBCO Laboratories has two new **serums** that it says provide an economical and more readily available alternative to fetal bovine serum (*Reader Service No. 114*). GIBCO's supplemented calf serum and calf supreme serum contain defined growth-promoting additives, including vitamins, trace elements, transferrin, insulin, and proprietary growth factors. GIBCO recommends its supplemented calf serum for researchers who experience lot-to-lot variability when using calf serum, and those for whom a high-performance basal calf serum would suffice. The company reports developmental studies indicate that a five per cent supplementation of calf serum in basal medium yields superior results over calf serum, and approaches results with fetal bovine serum. GIBCO says a five per cent supplementation of calf supreme serum in basal medium yields results comparable to fetal bovine serum.

A new version of the Sequenase **DNA polymerase** from United States Biochemical Corporation is now available (*Reader Service No. 115*). Version 2.0 of Sequenase was developed by *in vitro* genetic manipulation of the gene for T7 DNA polymerase which eliminated the exonuclease activity of the enzyme, giving it properties suitable for dideoxy DNA sequencing. US Biochemical says the new version retains all the advantages of the original Sequenase — low gel backgrounds, uniform band intensities, resolution of 600 or more bases, and the ability to use d1TP to remove gel compressions — but also demonstrates increased processivity and reduced pausing at regions of secondary structure. Sequenase Version 2.0 is available in 200- and 1,000-unit packages costing \$90 and \$355 (US), respectively, or as part of a kit with sufficient reagents for 100 sequences for \$185 (US).

Microscopy

International Scientific Instruments has just announced the availability of a **scanning electron microscope** that can be used to examine wet, unprepared samples (*Reader Service No. 116*). ISI's Wet-SEM



Wet or dry, International Scientific Instruments' Wet-SEM can handle any sample.

eliminates the need for sample preparation, such as coating, fixing, dehydrating or critical-point drying. Because the microscope operates in a low-vacuum environment, the \$72,500 (US) Wet-SEM can analyse moist specimens or even liquids. The uses of the Wet-SEM span a broad range of applications: it can be used with an optional EDX detector for elemental analysis, or can be switched to a high-vacuum setting for 6-nm resolution. Special features include an automatic focus system and an auto stigmator.

The Electron Microscope Supplies Division of EMCORP has a new **carbon-coater** that it says can deposit a thin film of carbon on specimens for electron microscopy in under two minutes (*Reader Service No. 117*). EMCORP's \$3,650 (US) Hummer CDS II comes ready to use right out of the box. It can coat specimens of up to 3.25 inches in diameter, or 12 microscope stubs in a single procedure. Its two-stage direct-drive rotary vacuum pump delivers at least 20 mtorr absolute pressure, and its manual bleed valve introduces argon for processing at 80–100

mtorr. The coater automatically vents once the main power switch is turned off to prevent backstreaming. The Hummer CDS II uses a pre-sharpened dual carbon rod system that EMCORP says yields consistent coating. Particle turbulence in the chamber is prevented by a regulated vent system.

Leitz Instruments Ltd is now supplying the Narshige range of **micromanipulators and ancillary equipment** with its Labovert and Fluovert inverted microscopes, and its Wild stereomicroscopes (*Reader Service No. 116*). Narshige's pipette puller, microforge, manual and automatic microinjectors, and micro beveller can now be purchased along with Leitz microscopes as a complete micromanipulation workstation which can be configured to suit a wide variety of applications.

Zeiss is bringing the line of **rotary microtomes** made by the German company Microm to the US market (*Reader Service No. 119*). Microm microtomes come in designs for routine paraffin histology, advanced plastics histology and industrial microtomy. Zeiss says the Microm microtomes have backlash-free and maintenance-free roller bearing guideways, enabling them to achieve high precision and stability. Specimen slices between 0.5 µm and 60 µm in thickness are routinely achieved, according to the company. □



Carbon coating of EM samples in under two minutes by EMCORP.

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

What prospects for the older graduate?

Richard Pearson

Employers have been slow to adjust to the changes in graduate supply. Soon, with decreasing numbers of school leavers, they may have to.

LAST year was the year of the school leaver, or rather concern about the lack of them in the 1990s: as the number of school leavers in the United Kingdom will have declined by 35 per cent over the decade to 1994. With the economy booming and unemployment falling, a shortage of young people will affect employers, trainers and higher education alike (*Nature* 335, 100; 1988). In the case of higher education the hope is that more mature students can be found who, along with other under-represented groups such as women, ethnic minorities and those from the lower social classes, will, along with a rising participation rate among young people, fill the empty places in higher education and boost the numbers graduating to meet the needs of the economy.

Since the late 1970s there has been growing interest in improving mature students' access to higher education, with more than 400 special access courses being launched to help mature students who lack the standard qualifications for access to higher education. At the same time most institutions now offer applicants aged over 21 selection criteria that do not rely on the normal academic qualifications. Just over half the mature entrants to full-time degree courses lack the traditional entry qualifications, rising to 75 per cent for those entering part-time courses.

Mature numbers

During the past decade the number of mature entrants in higher education has increased by over 50 per cent to total over 186,000 in 1986, driven mainly by the growing number of women and those on part-time courses, the number of men entering full-time university courses rising by only 10 per cent. The overall proportion (aged 21 or over at the time of entry) has now reached over 11 per cent in the universities and 34 per cent in the polytechnics and colleges. The majority of mature students are, however, studying at sub-degree level and on a part-time basis. These older students also have rather different subject preferences to their younger counterparts, with a much higher representation in the humanities and the social sciences, and lower representation in the sciences, technology and engineering.

In addition there is the Open University which was established to meet the needs of the mature, part-time student, many of whom are already in employment. Some 8,000 students graduate each year from

the Open University, and of new entrants over 60 per cent are aged 30 or over.

It is worth noting that while in the United Kingdom the mature first degree graduate is still relatively rare, the majority being aged 21–22, this contrasts with the pattern in mainland Europe where courses take longer and many graduates only qualify in their mid to late 20s.

In trying to achieve a further increase in the number of mature entrants, institutions will be facing growing competition from employers for the good mature entrant as both they and higher education seek to make up for the shortfall of school leavers. One important determinant of who wins this battle will be students' responses to loans and the way they trade cash in the hand for potentially higher earnings later in life. One concern, expressed even before the loans proposals were introduced, is the financial barrier facing students taking access courses who were not eligible for mandatory grants. A recent survey of students suggested that half of them would not have started their access courses if the grant had not been available.

What then are the prospects for mature graduates, and will they benefit alongside the traditional young graduate in the booming graduate labour market of the 1990s (*Nature* 334, 90; 1988)? Looking back to the mid-1980s, when the graduate market was starting to take off, mature graduates were slightly less likely to be in employment and also less likely to be unemployed, with a higher percentage going into further education and training or not being available to work. The latter figure is rather higher still for mature women than for mature men.

When due allowance is made for subject differences mature graduates had rather better employment rates in education, with little difference being apparent in the case of engineering. In the case of the sciences, administration and business studies, and humanities, they fared rather worse, with both lower employment and higher unemployment rates.

How then are mature entrants treated in the labour market? One problem they face is that many employers, especially in the private sector, structure their recruitment, training and career development programme around the 21 year olds, seeking to get them into management positions by their late 20s. A recent survey has, however, shown that while half of the

respondents say they treat all their applicants equally, one in three expressed reservations about applications from mature graduates, with one in five specifying age limits, normally 30 or under, in their recruitment literature. Teaching is a notable exception where positive efforts are being made to attract mature entrants.

Most welcome

More generally the most welcomed were those applicants who had a relevant degree and relevant work experiences, although there was still an age threshold of about 30 after which entry could become more difficult. Much depends on the relevance of the work experience and the job applied for, it being much harder to enter a general training scheme than entering a direct appointment. For those with a non-relevant degree the opportunities were more restricted, reflecting in part the more limited opportunities open to this group of conventional graduates. In general, difficulties increased with age and with the decreasing relevance of qualifications.

The main perceived disadvantages of the mature graduate were their apparent poor mobility, poor adaptability, too high salary expectations, difficulties fitting in with younger peers and not fitting into promotion structures. Given that many companies do not recruit them, it must be assumed that they have a strongly stereotyped picture of them. On the positive side their maturity, stability, commitment and better social and interpersonal skills were often highlighted. Another positive sign for those in their mid-20s is recruiters' growing interest in 'second bounce' graduates, those who re-enter the job market quickly after a mistaken first choice.

With the market expected to tighten further into the 1990s only a minority of recruiters when surveyed say they are looking to adjust their policies, despite the fact that mature graduates will be an increasing proportion of those graduating in the future. As with school leavers, most recruiters seem set to carry on with policies appropriate to a time of abundance in the labour market rather than one of stiffening competition. For the mature graduate, economic conditions and skill shortages are going to be the fastest way of breaking down traditional stereotypes. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Falmer, Brighton BN1 9RF, UK.